TECH INDIA PUBLICATION SERIES

INTRODUCING PROBABILITY & STATISTICS

A Problem-Oriented Uncluttered Friendly Approach
With more than 900 fully Worked-out Distinct Problems

Dr. Bansi Lal
Dr. Sanjay Arora
Sudha Arora

SOME FORMULAS

$$\begin{cases} P(B) = P(B \mid A_1)\,P(A_1) + P(B \mid A_2)\,P(A_2) + \ldots + P(B \mid A_n)\,P(A_n) & \text{[Multi-Stage } p\text{-Rule]} \\ P(B \mid C) = \Sigma P(B \mid A_i C)\,P(A_i \mid C), \quad i = 1, 2, \ldots n; & \text{[Multistage Conditional } p\text{-rule]} \\ E(X) = P(A_1)\,E(X \mid A_1) + P(A_2)\,E(X \mid A_2) + \ldots + P(A_n)\,E(X \mid A_n) & \text{[Multistage } E\text{-rule]} \end{cases}$$

$$P\{B_i \mid A\} = \frac{P(A \mid B_i)\,.\,P(B_i)}{\Sigma P(A \mid B_i)\,P(B_i)}$$ [Bayes' Reversal Rule]

$P(A_1 A_2 \ldots A_n) \geq P(A_1) + P(A_2) + \ldots + P(A_n) - (n-1)$ [Bon-ferroni inequality]

$P(A_1 \cup \ldots \cup A_n) \leq P(A_1) + \ldots + P(A_n)$ [Boole's inequality]

$E[(X-c)^k] = k\int_c^\infty (x-c)^{k-1}\,[1 - F(x)]\,dx - k\int_{-\infty}^c (x-c)^{k-1}\,F(x)\,dx.$

$P\{|X - \mu| \geq c\} \leq \sigma^2/c^2$, or $P\{|X-\mu| < c\} \geq 1 - (\sigma^2/c^2)$ [Chebyshev inequality].

$P\{|X - b| \geq a\} \leq E(X-b)^m / a^m.$ [Non-central Chebyshev inequality]

Let X and Y be *independent*. Let $S = X + Y$, $D = Y - X$, $P = XY$, $Q = Y/X$. Then

$$\boxed{\begin{aligned} g_S(z) &= \int_{-\infty}^\infty f(x, z-x)\,dx = \int_{-\infty}^\infty f(z-y, y)\,dy = \int_{-\infty}^\infty f_1(x)\,f_2(z-x)\,dx = \int_{-\infty}^\infty f_1(z-y)\,f_2(y)\,dy \\ g_D(z) &= \int_{-\infty}^\infty f(x, z+x)\,dx = \int_{-\infty}^\infty f(z+y, y)\,dy = \int_{-\infty}^\infty f_1(x)\,f_2(z+x)\,dx = \int_{-\infty}^\infty f_1(z+y)\,f_2(y)\,dy \\ g_P(z) &= \int_{-\infty}^\infty f\left(x, \frac{z}{x}\right)\frac{dx}{|x|} = \int_{-\infty}^\infty f\left(\frac{z}{y}, y\right)\frac{dy}{|y|} = \int_{-\infty}^\infty f_1(x)\,f_2\left(\frac{z}{x}\right)\frac{dx}{|x|} = \int_{-\infty}^\infty f_1\left(\frac{y}{z}\right)f_2(y)\,\frac{dy}{|y|}. \\ g_Q(z) &= \int_{-\infty}^\infty f(x, zx)\,|x|\,dx = \int_{-\infty}^\infty f\left(\frac{y}{z}, y\right)\frac{|y|}{z^2}\,dy = \int_{-\infty}^\infty f_1(x)\,f_2(zx)\,|x|\,dx = \int_{-\infty}^\infty f_1\left(\frac{y}{z}\right)f_2(y)\,\frac{|y|}{z^2}\,dy. \end{aligned}}$$

$P(X = a) = F(a) - F(a-0)$ [Jump]. $E[\varphi(X)] \geq \varphi[E(X)]$ [Jensen's inequality]

$E[g(X)] = \Sigma_x f(x)\,g(x)$ [Kernel-summation], $E[g(X)] = E\{E[g(X) \mid Y]\}$, $E(X) = E[E(X \mid Y)]$. [Double-E Rule]

$\lim_{n\to\infty} P\left\{\left|\frac{S_n}{n} - E\left(\frac{S_n}{n}\right)\right| < \varepsilon\right\} = 1$ [WLLN] $\qquad \frac{S_n - E(S_n)}{\sigma(S_n)} \xrightarrow{d} N(0,1)$ [CLT]

$$\begin{cases} \text{Bivariate Normal Dist. on : BVN}(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2; \rho): \quad f(x,y) = (2\pi\sigma_1\sigma_2\sqrt{1-\rho^2})^{-1}\exp(-\tfrac{1}{2}Q), \\ \text{where } Q = \left[\left(\frac{x-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{x-\mu_1}{\sigma_1}\right)\left(\frac{y-\mu_2}{\sigma_2}\right) + \left(\frac{y-\mu_2}{\sigma_2}\right)^2\right]\Big/(1-\rho^2), \ -\infty < x, y < \infty \\ M(t_1, t_2) = \exp[\mu_1 t_1 + \mu_2 t_2 + \tfrac{1}{2}(\sigma_1^2 t_1^2 + \sigma_2^2 t_2^2 + 2\rho\sigma_1\sigma_2 t_1 t_2]. \\ X \sim N(\mu_1, \sigma_1^2);\ Y \sim N(\mu_2, \sigma_2^2); \quad (Y \mid x) \sim N[\mu_2 + (\rho\sigma_2/\sigma_1)(x - \mu_1),\ \sigma_2^2(1-\rho^2)]. \end{cases}$$

$\chi^2_{(n)} = Z_1^2 + Z_2^2 + \ldots + Z_n^2$, $\quad Z_j$ are i.i.d. $N(0, 1)$.

For $N(\mu, \sigma^2)$ Distn.; $\overline{X}$ and $\hat{S}^2$ are independent; $(nS^2/\sigma^2) \sim \chi^2_{(n)}$; $\quad (\overline{X} - \mu)/(\hat{S}/\sqrt{n}) \sim t_{(n-1)}$
$N(0,1)/(\chi^2/n)^{1/2} \sim t_n$; $\quad [\chi^2_m/m]/[\chi^2_n/n] \sim F(m,n)$.

$F(m,n)$: $f(x) = \dfrac{(m/n)^{m/2}}{B(\frac{1}{2}m, \frac{1}{2}n)}\dfrac{x^{(m/2)-1}}{[1 + (m/n)x]^{(m+n)/2}}, x > 0.\quad t_{(n)} : f(y) = \dfrac{1}{\sqrt{nB(\frac{1}{2}, \frac{1}{2}n)}}\left(1 + \dfrac{y^2}{n}\right)^{-(n+1)/2}, |y| < \infty$

F-dist$^{\mathbf{n}}$: $E(X) = n/(n-2)$, $\quad$ Var $(X) = n^2(2n + 2m - 4)/m(n-2)^2(n-4),\ n > 4$

t_n-dist$^{\mathbf{n}}$: $E(Y) = 0 (n > 1)$, $\quad$ Var$(Y) = n/(n-2)$, if $n > 2$; mgf does not exist.

$t_{(1)} =$ Chy $(0, 1)$, $\quad$ gam(a, λ)/gam$(b, \lambda) \sim B_{II}(a, b)$; $N(0,1)/N(0,1) \sim$ Chy$(0, 1)$

INTRODUCING
PROBABILITY & STATISTICS

A Problem-Oriented Uncluttered Friendly Approach
With more than 900 fully worked-out distinct problems

INTRODUCING PROBABILITY & STATISTICS

A Problem-Oriented Uncluttered Friendly Approach
With more than 900 fully worked-out distinct problems

Dr. Bansi Lal
Ex. Deptt. of Mathematics
K.M.College, University of Delhi (India)
Formerly, Visiting Professor of Mathematics
University of Iowa, Iowa City, U.S.A.

Dr. Sanjay Arora
Ex. Deptt. of Mathematics
Sri. Venkateswara College, Univ. of Delhi
Howard University,
Washington, D.C., USA

Sudha Arora
Deptt. of Mathematics
Deen Dyal Upadhyaya College
University of Delhi
Delhi, India

Satya Prakashan
Incorporating : TECH INDIA PUBLICATIONS
NEW DELHI-110005

Published by :

Smt. Sumitra Handa

Prop. ***Satya Prakashan,***

16/7698, New Market, (Near Liberty Cinema) New Rohtak Road, Karol Bagh, New Delhi-110 005.

PHONE : 28713653, 28715679, **MOBILE :** 9212081410, 9810124893, 9810061181 **FAX :** 28713593

E-MAIL : ◆ satyaprakashan@gmail.com ◆ deepakhandarph@gmail.com

WEBSITE : www.satyaprakashan.com

First Edition	:	1994
Reprint Edition	:	2000
Second Edition	:	2006
Reprint Edition	:	2009
Reprint Edition	:	2012
Reprint Edition	:	2014-15
Reprint Edition	:	2018

ISBN : 81-7684-339-3

Price : **Rs. 350.00**

Editor : Deepak Handa

Production Incharge : Ajay Handa

Laser Set by : Aisa Enterprises

Cover Design by : R.P. Computers PHONE 9810124893

Printed at : Sumitra Enterprises : Karol Bagh, New Delhi PHONE 9810061181

Every weakness has found its defenders ; but stupidity hasn't (Prim Levi)

Preface to the Second Edition

This text is designed as an introduction to elements of probability theory and Mathematical Statistics, aiming to reach the mid-high, but beginning at the lowest level. The approach at places is thus heuristic and less rigorous, using elementary knowledge of mathematics. The attempt is to develop in the student an intuitive as well as mathematical feel for the subject, which enables the reader to think scientifically. Not all topics of the subject are included, which are otherwise presented by the authors in their Text on Mathematical Statistics (1989). But topics presented are in greater detail and profusely illustrated.

Additions/Improvements to this Edition

- A very brief and naive knowledge of random variables and their expected values is included in Chapter 1 (§1-42 & §1-43). This has resulted in shorter solutions to some problems by using mean values of binomial and Hyp-geom variates. Besides, it has provided notational advantage.
- Some Combinatorial solutions are replaced by routine solutions which are based on integer-solutions to the equation $x_1 + x_2 + ... + x_n = m$ where $x_j > 0$ or $x_j \geq 0$, $1 \leq j \leq n$. The proof of this result as well as several other results (with or without proofs) are presented in section "Some Useful Formulas" pages 819-829. This shall secure firmness for math-feel.
- Useful comments on Surprise and Uncertainty (§1-9) will incite interest for Modern Information Theory.
- Multistage p-Rule, Multistage Conditional p-Rule (§2-10) and Multistage E-Rule (§5-41), frequently called *conditioning processes*, have an infinite potential to handle unlimited number of situations in the simplest possible manner. This extends the Frontiers of knowledge and we have exhibited it frequently.
- Indicators play an excellent shortening-device role in very laborious evaluations in Probability Theory. Their use calls for more intelligent manupulations, frequently asks for a sharp imagination. Their use in §1-53 and §5-43 is remarkably demonstrated.
- To smash the myth that Conditional Expectation (Chapter 7) is more involved, we include more, Solved Problems on this valuable concept. This must expel any pseudo tensions, if any.

- Characteristic Functions for integer-valued periodic functions (§9-40). A direct evaluation of the Ch. function of Cauchy distribution (§9-70).
- Structural similarity between bin (n, p) and Hyp-geom (N, M ; k) in §15-21.
- For solutions to Numericals on Normal Distribution [Chap 16], both types of Statistical Tables : $P(-\infty < Z \le z)$ and $P\{0 \le Z \le z)$ are provided for ease and bent of mind of the individual solver.
- Proof of Stirling's approximation to n! by using CLT [§17-13]. Our proof is more elementary than the existing ones.
- Geometric presentation of Rejection Method [§18-31]. This shall inculcate Simulation of Variates more smoothly.
- This Edition contains more than 900 fully worked-out distinct problems. For their rapid coverage, 542 problems are starred and their solutions are provided at the end of the Text [pp. 659-818]. This shall provide flexibility, encourage self-attempts and the speedy completion of the course contents.
- Each chapter or its Appendix, starts and ends with a witty quote. Ponder over them for amusement and even for learning.

Warmest thanks are due to Prof. Surjit Singh Khurana (Iowa University, USA), Prof. Kanwar Sen (University of Delhi) and Dr. K.P. Chinda (Principal Keshav M.V. Univ. of Delhi) for their helpful discussions. I am also thankful to the publishers for their diligence.

I owe a gratitude to Prof. Jiwan Lal Tiwari, who taught me mathematics for my B.A. (Hons) Course in session 1951-1953. What I learnt in mathematics is a lot due to his efforts. I am extremely indepted to him and shall remain ever indebted for his kindness, courtsy and helpful attitude.

1-Vaishali, Pitampura, Delhi-88 **BANSI LAL**
Phone (R) 27317618
1st December 2005

Preface to the First Edition

This book provides an extensively detailed and reasonably rigorous introduction to Basics of elementary parts of Probability Theory with emphasis on Probability Models and some Statistics. The material is based on the thorough revised version of several Chapters from "New Mathematical Statistics" (1989) and the target space is the new course structured for Mathematics, Honours of Delhi University, although its periphery is all Indian Universities.

The major goal of this text is to help the reader digest the fundamental ingradients of Probability and its allied concepts without much pain and effort. To achieve this objective, clear statements of pertinent Definitions, Principles, Theorems and Comments, together with 879 numbered solved problems are included. There are a large number of exercises to test one's grasp of the subject. These solved examples, ranging from insultingly easy to

slightly difficult, mostly adopted from Indian Universities examination papers and Public Service Competition Examinations, serve to illustrate and amplify the basic structure of Probability concepts. Besides their examination values, these solved problems bring into sharp focus such fine points without which the students continually feel themselves on unsafe ground, and provide the repetition of basic principles which are vital to effective learning. The geometric-analytic duality which is inherent in "Introducing Probability & Statistics" pervades throughout the text.

The text begins with the prime concept of Probability Set function and culminates with extensive study of probability to various probability models. To review completely the subject matter of each chapter, a good number of supplementary exercises are included in each major section of the concerned chapter. Some exercises embody counter examples, some probe the extensions to existing ones, while most of them are routine : getting-your-feet-wet type. Answers to most of the exercises are supplied or embedded in their structure.

A glance over the table of contents will indicate to the reader some idea of material covered in this text book. Compared with most books, much elementary and lighter material, besides some slightly advanced stuff, consistent with Indian Back requirements, is included in an instantly digestible form. The study of Probability & Statistics is slanted towards elementary mathematics rather than to commercial sciences. The reader will discover in this text, that several problems worked-out elsewhere, have much shorter and amazingly much neater solutions. The fine details provided to complicated situations will relieve the readers of their tensions. The relatively advanced material is generally, separated out from the main stuff so that choice of adoption is easily made. Appendices A, B, C, D, E and F as well as Chapters 7, 9, 10, 11 and 17 are pointers in this direction. Quite a lot of stuff appears for the first time in simple digestible form, e.g. §16-61 and §18-40 and so on. Many solutions shall strike novelty and better mathematization, mainly due to using Multi-Stage Rules, Generating functions and integration simultaneously involving indicators e.g. §5-90.

In conclusion, we remark that it has been our endeavour to design a text-book that will help the Readers at all levels to have a grasp of fundamental concepts of Probability & Statistics that will be found useful in Mathematical Analysis and in all Pure and Applied Sciences, such as Physics, Chemistry, Computer Science, Biology, Education, etc. In quest of this objective, we have consulted a large number of books and a brief Bibliography at the end speaks of our gratitude to their Authors & Publishers.

During the preparation of the MSS, we were guided by the advice and suggestions of Prof. D. L. Jain, Head : Dept. of Mathematics and formerly Dean : Faculty of Mathematical Sciences, University of Delhi, Prof Surjit Singh Khurana, Mathematics Department, University of Iowa, Iowa City, U.S.A.; Dr. James A Donaldson, Mathematics Department, Howard University, Washington D.C., U.S.A. and Dr. Y. P. Sabharwal, Department of Mathematics & Statistics, Ramjas College, University of Delhi. To them, we owe much more than words at our command can express.

We are also obliged to Dr. Ajay K. Arora (Hans Raj College, University of Delhi) who has rendered the valuable service of reading the cumbersome proofs during the printing

process and also prepared the Subject Index to this text.

In the production process, Miss Indu Nair & V K Nair at Kamala Compu Compose, did a marvellous job and we are grateful to them for their courtsy, sincere and hard work extended by them. The Publishers : Deepak Handa & Ajay Handa, on their part have shown full appreciation of the task presented to them and deserve our gratitude.

Finally, we put on record that without the encouragement and support of Mrs. Sarla Rani, this book would have remained only a dream.

The authors will appreciate receiving the intimation of misprints and misconcepts and shall duly acknowledge such intimations with gratitude.

1 Vaishali, Pitampura
Delhi 110 034

SANJAY ARORA
BANSI LAL

Those who stand for nothing, fall for anything. (Alex Hamilton)

Dedicated to my daughter-in-law

Santosh (Raju)

BANSI LAL

Dedicated to my mother

Sarla Arora

SANJAY ARORA

Dedicated to my daughters

Arunima & Aarushi

SUDHA ARORA

Before God, we are all equally wise – and equally foolish (Albert Einstein)

Partial Contents

Most paid jobs absorb and degrade the mind (Aristotle)

It is not certain that everything is uncertain. *(Blaise Pascal)*

Elementary Probability. Basic Concepts

1

1-10. *Introduction*

The origin of probability theory stemmed from the analysis of certain games of chance popular in the seventeenth century. It has since found applications in many branches of science and engineering and this extensive application makes it a valuable study. Probability Theory is, as a matter of fact, a study of random or non-deterministic experiments and is helpful in investigating the regular features of these random experiments.

By a *random experiment*, we mean an experiment in which :

(i) All the outcomes of the experiment are known in advance.

(ii) Any performance of the experiment results in an outcome that is not known in advance.

(iii) The experiment can be repeated under identical conditions.

The experiment in which we can completely predict the outcome is known as *deterministic experiment.* For example, the formula $x = \frac{1}{2}gt^2$ gives the position of the particle falling under gravity from a position of rest, at any time t. However, in random experiments, the situation is different. Consider the following situations :

1. A fair coin is tossed. We assume that the coin does not stand on its edge. On any toss, it is known that either head or tail will appear but it is not known exactly what the outcome will be. Obviously, the coin may be tossed any number of times.

2. A fair die is rolled. It is known that any of the six possible outcomes will occur but it is not known what exactly the outcome will be. The die may be rolled any number of times.

In the present chapter, we shall first discuss the **Classical** (mathematical) or a **priori** approach to Probability Theory, which will be followed by the Statistical (frequency or posterior) and the **axiomatic** approaches. The relative merits of three approaches shall also be examined. Before doing so, we settle some terminology and notations.

1-11. *Some Notions and Notations*

1. ***Sample Space.*** Given any random experiment, the set of all possible outcomes of this experiment is called a *sample space* or *outcome space.* This set is denoted by Ω (omega) or S and the individual outcomes are generally denoted by ω or s.

Some Illustrations :

E_1. Toss a coin once : $S_1 = \{T, H\}$ (T = tail, H = head)

E_2. Toss a coin twice : $S_2 = \{(T, T), (T, H), (H, T), (H, H)\}$

E_3. Roll a single Die : $S_3 = \{1, 2, 3, 4, 5, 6\}$ (Number of eyes)

E_4. Roll of a pair of Dice : $S_4 = \{(x, y) : x = 1, 2, ..., 6 ; y = 1, 2, ..., 6\}$

E_5. Toss a coin and roll a Die : $S_5 = \{(H, 1), (H, 2), ..., (H, 6), (T, 1), (T, 2), ..., (T, 6)\}$

Remarks. The sample space of a random experiment need not be unique. The particular sample space considered is a matter of convenience and expediency. For example, the sample space for the experiment E_2 may well be represented by $S_2' = \{0, 1, 2\}$, where 0 means no head, 1 means one head and 2 means two heads.

2. *Events.* the subsets of the sample space are called *Events*.

(i) An event E is called *elementary event* if it consists of only one element, i.e. when E is singleton.

(ii) An event which is not elementary is called *compound event*.

We say that the event A occurs if the outcome of the experiment corresponds to the elements of A.

3. *Sure Event.* Since S itself is a subset of S, so S is an event. Also $s \in S$ for all s, so the event S occurs at each observation, it follows that the universal set (sample space) S is the *sure (certain)* event.

4. *Impossible Event.* The empty set $\varnothing$ is a subset of S (the sample space), hence $\varnothing$ is an event, as per our definition of event. Also $s \notin \varnothing$ for all $s \in S$, so that (empty set) event $\varnothing$ does not occur at any observation. As such $\varnothing$ is called the *null* or *impossible event*.

5. *Event Space.* Let S denote the sample space of a random experiment. Then the set of all events including S and $\varnothing$ is called an *event space*; it is usually denoted by ε.

The set ε is a very big set as compared with S. In fact, if S contains n elements, (i.e. S is finite), then ε will contain 2^n elements. In axiomatic treatment of Probability, we shall show how to reduce the event space.

6. *Mutually Exclusive Events.* Events A and B are called **disjoint** or *mutually exclusive* (or incompatible) if $A \cap B = \varnothing$. That is, A and B are alternate possible outcomes of the same trial.

Notes. (i) Two distinct elementary events are always disjoint i.e. $\{a\} \cap \{b\} = \varnothing$.

(ii) Every compound event can be uniquely represented as the union of a set of elementary events, i.e.

$$A = \{s_1\} \uplus \{s_2\} \uplus \ldots \uplus \{s_n\} \qquad [\text{See below for } \uplus]$$

7. Specific Terminology. Let $A, B, C, \ldots$ be possible events. We shall follow as under :

Notation	*Meaning*
$A \cup B$	events A or B or both [at least one of two events]
$A \cap B$ or AB	events A and B [both occurring together,
$A \uplus B$	event A or B but not both [i.e. with $AB = \varnothing$]
$A \subset B$	A implies B [A is a sub-event of B]
$A = B$	events A and B are identical
A^c or $\bar{A}$ or A'	event opposed to A [complementary to A]

$A - B = A \cap \bar{B}$	event A occurs but B does not occur
$A \mid B$	event A conditional by the occurrence of event B.
$A \Delta B = A\bar{B} \uplus \bar{A}B$.	either A or B but not both.

Important note :

(a) We shall frequently write AB instead of $A \cap B$.

(b) $A \uplus B$ means $A \cup B$ together with $A \cap B = \varnothing$.

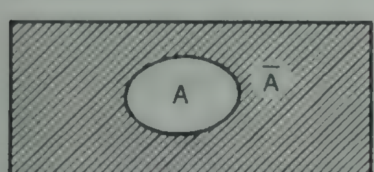
(c) $A \uplus B$ is an out-dated and ill-conceived expression for $A \cup B$ or $A \uplus B$.

(d) Extensions to more than two events and their corresponding meanings are obvious, e.g.

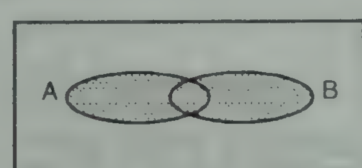
$$A \cup B \cup C,\ A \cap B \cap C,\ A \cap (B \cup C),\ \left\{A\bar{B}\bar{C} \cup \bar{A}B\bar{C} \uplus \bar{A}\bar{B}C\right\},\ \cup_i A_i,\ i - 1, 2, \ldots$$

Associative, distributive and complement laws from Set Theory are frequently employed.

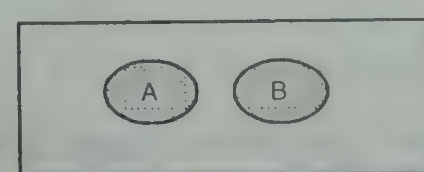
8. *Venn Diagrams.* The diagrams, depicting the unions and intersections of various events are called *Venn diagrams* or *Euler diagrams*. Some simple combinations of the events (shaded), and their expressions are demonstrated below; the rectangle is denoting the sample space S.



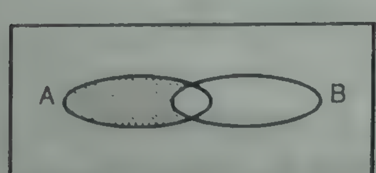
(i) $\bar{A}$



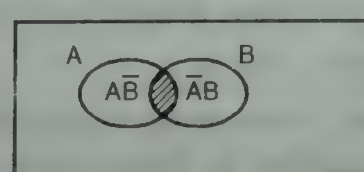
(ii) $A \dot{\cup} B$



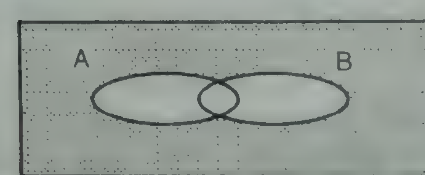
(iii) $A \uplus B$



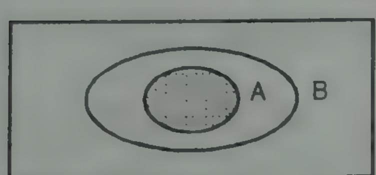
(iv) $A\bar{B}$



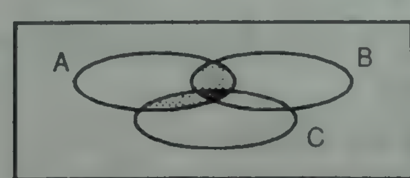
(v) $A \cap B$



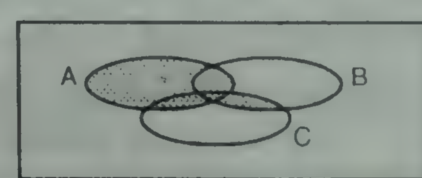
(vi) $\bar{A} \cap \bar{B} = \overline{(A \cup B)}$



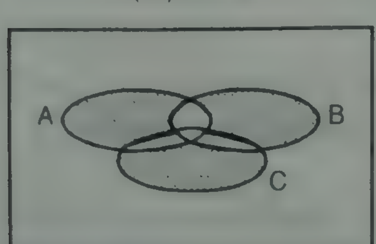
(vii) $A \subset B$



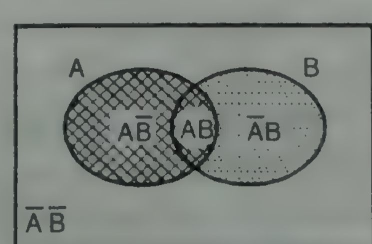
(viii) $AB \cup AC = (A \cup B) \cap (A \cup C)$



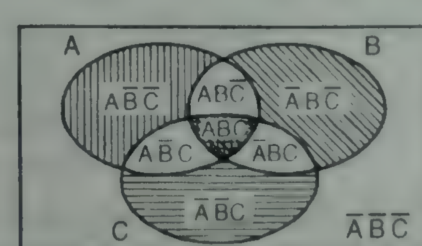
(ix) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$



(x) $A \uplus B \uplus C$



(xi) $A \cup B = A\bar{B} \uplus AB \uplus \bar{A}B$



(xii) $A \cup B \cup C$

A figure of particular utility is the **Euler diagram** (xi) for two possible events A and B where events $AB, A\bar{B}, \bar{A}B,\ \bar{A}\bar{B}$ are demonstrated and some facts like :

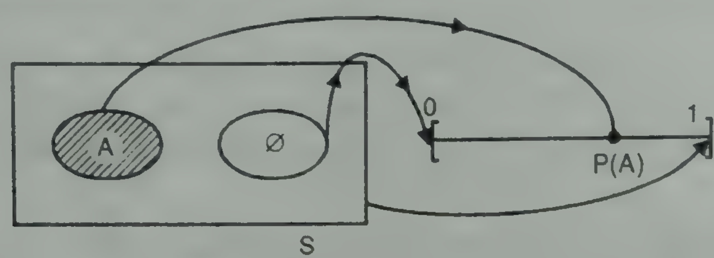
$A \cup B = A \uplus \bar{A}B$, $A = A\bar{B} \cup AB$, $\bar{A}\bar{B} = S - (A \cup B)$, $AB \subset A$, etc. are visually shown.

9. *Partition or Exhaustive Set of Events.* A set $\{A_i\}$ of disjoint events, $(A_i \cap A_j = \varnothing, \forall_i \neq j$ is called *exhaustive set of events* (or Partition of S) if $\biguplus_i A_i = S$. This means that every point of the sample space belongs to one and only one of the events A_i.

1-20. *Classical (or Laplace) Definition of a Probability Function*

Let S be an outcome space (assumed to be finite) of a given random experiment. Let $n(S)$ be the total number of possible outcomes which are mutually exclusive (m.e.) and whose occurrences are equally likely. Let $A \in \mathbb{E}$ (event space) be an event with $n(A)$ possibilities. The probability of event A occurring, written $P(A)$, is defined by the number

$$P(A) = \frac{n(A)}{n(S)} = \frac{\text{Number of cases favourable to } A}{\text{Exhaustive number of cases in } S} \qquad \text{[A priori approach]} \quad ...(1)$$



Obviously $\qquad 0 \leq P(A) \leq 1$, since $0 \leq n(A) \leq n(S)$, $\qquad$ [Probability Scale].

1-30. *Two Basic Laws of Probability Theory*

1. *Addition Law for Two Events* :

If A and B are any two arbitary (possible) events in a sample space S, then

$$\underline{P(A \cup B) = P(A) + P(B) - P(AB)} \qquad ...(1)$$

Proof. Denote the number of cases favourable to the events A, B, $A \cup B$ and $A \cap B$ by $n(A)$, $n(B)$, $n(A \cup B)$ and $n(AB)$ respectively. By counting the points in various sets of S, we see that (See the Euler's diagram)

$$n(A) = n(AB) + n(AB), \quad n(B) = n(AB) + n(\bar{A}B), \qquad n(A \cup B) = n(AB) + n(AB) + n(\bar{A}B).$$

Eliminating $n(A\bar{B}), n(\bar{A}B)$ provides

$$n(A \cup B) = n(A) + n(B) - n(AB).$$

Divide this equation throughout by $N = n(S)$, to get

$$\frac{n(A \cup B)}{N} = \frac{n(A)}{N} + \frac{n(B)}{N} - \frac{n(AB)}{N}.$$

By Laplace (Classical) definition of mathematical probability, this equation is equivalent to

$$P(A \cup B) = P(A) + P(B) - P(AB).$$

Cor. If $A \cap B = \varnothing$, i.e. if the events are mutually exclusive, then there is no point favourable to the simultaneous occurrence of A and B so that $P(AB) = 0$. Equation (1) then reduces to

$$\underline{P(A \uplus B) = P(A) + P(B)}.$$

Extension to three events. For three events A, B, C

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - [P(AB) + P(BC) + P(AC)] + P(ABC).$$

Proof. Replace B by $B \cup C$ in (1) to obtain

$$P(A \cup B \cup C) = P(A) + P(B \cup C) - P[A \cap (B \cup C)]$$
$$= P(A) + P(B \cup C) - P(AB \cup AC), \qquad \text{[Distributive Law]} \quad ...(1)$$

Apply (1) to (i); use $AB \cap AC = ABC$, to obtain

$$P(A \cup B \cup C) = P(A) + [P(B) + P(C) - P(BC)] - [P(AB) + P(AC) - P(ABC)]$$
$$= P(A) + P(B) + P(C) - [P(AB) + P(BC) + P(AC)] + P(ABC). \quad ...(2)$$

Comments. Write $S_1 = P(A) + P(B) + P(C)$, $S_2 = P(AB) + P(BC) + P(AC)$, $S_3 = P(ABC)$.

$$\therefore \quad P(A \cup B \cup C) = S_1 - S_2 + S_3. \quad ...(2)$$

2. *Theorem of Compound Probability or Product Rule :*

If A and B are any two possible events of a sample space S, then

$$\underline{P(AB) = P(A)\, P(B \mid A) = P(B)\, P(A \mid B)} \quad ...(1)$$

Events	A	$\bar{A}$	*Totals*
B	n_{11}	n_{12}	$n_{11} + n_{12}$
$\bar{B}$	n_{21}	n_{22}	$n_{21} + n_{22}$
Totals	$n_{11} + n_{21}$	$n_{12} + n_{22}$	N

Proof. Denote the number of cases favourable to the events AB, $\bar{A}B, A\bar{B}$ and $\bar{A}\,\bar{B}$ by $n_{11}, n_{12}, n_{21}, n_{22}$ respectively (as shown in the above Table). Let N be the total numbers of cases in the sample space. By Laplace (classical) definition of the mathematical probability,

$$P(A) = \frac{n_{11} + n_{21}}{N}, \quad P(B) = \frac{n_{11} + n_{12}}{n}, \quad P(AB) = \frac{n_{11}}{N}, \quad P(A \mid B) = \frac{n_{11}}{n_{11} + n_{12}}. \quad ...(i)$$

Here $P(A \mid B)$ = probability of event A relative to the space of event B. [This is the obvious classical Def. of conditional probability]

$$\text{Obviously,} \quad P(A \mid B) = \frac{n_{11}}{n_{11} + n_{12}} = \frac{n_{11}/N}{(n_{11} + n_{12})/N} = \frac{P(AB)}{P(B)}, \quad \text{[by (i)]} \qquad [P(B) \neq 0].$$

We rewrite the above equation to obtain $P(AB) = P(B)\, P(A \mid B)$.

The other result follows by interchanging A and B.

Extension to Three Events. For the possible events A, B, C

$$\underline{P(ABC) = P(A)\, P(B \mid A)\, P(C \mid AB).}$$

Proof.
$$P(ABC) = P[(AB)C] = P(AB)\, P(C \mid AB)$$
$$= P(A)\, P(B \mid A)\,.\, P(C \mid AB). \qquad \text{[by (1)]}$$

Definition. Events A and B are said to be *independent* if

$$P(A \mid B) = P(A), \qquad \text{or} \qquad P(B \mid A) = P(B).$$

Product rule for two and three independent events :

$$P(AB) = P(A)\, P(B), \qquad P(ABC) = P(A)\, P(B)\, P(C).$$

Remarks. Solutions to problems of Probability and Statistics require various manipulative skills, which are quoted for constant use in the Chapter zero of this book.

1-40. *Procedure for Solving Probability Problems*

1. Choose an appropriate sample space which includes as subsets all of the events A_i to be considered.
2. Identify all of those events A_k to which reasonable probabilities (ordinary or conditional) can be assigned and make these assignments.
3. Describe the event A for which $P(A)$ is to be computed.
4. Relate A to the other events A_i, which fix up A.
5. Apply the various laws of Probability and the definitions of allied nature, e.g. independence, conditional probability, disjointedness, etc. Compute $P(A)$ and examine if this makes sense.
6. If $P(A)$ is not precise, find the good approximation or useful bounds for $P(A)$.

Comments. In practice, some of the steps are mixed up together, and (6) is seldom needed.

1-41. *Worked-out Problems*

***Example* 1.** Three numbers are selected at random one after another and without replacement from m numbers 1, 2, ..., m. Find the probability that : (i) The first number drawn is the smallest and the second number the largest; (ii) The first number is smaller than the second number.

Solution. To each triplet of numbers out of $\{1, 2, ..., m\}$ there correspond six permutations (ordered triplets). Let us suppose that three numbers withdrawn are a, b, c and that $a < b < c$. The possible sequences of withdrawals are :

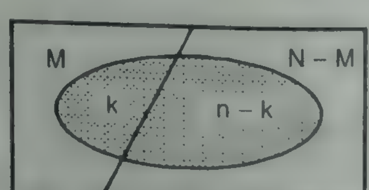
$$S = \{abc, acb, bca, bac, cab, cba\} \text{ : six possible outcomes.}$$

(i) The event of interest $A = \{acb\}$, has the probability $P(A) = 1/6$.

(ii) The event of interest $B = \{abc, acb, bca\}$, has the probability $P(B) = 1/2$.

***Example* 2.** An urn contains N balls numbered 1, 2, 3, ..., N. Find the chance that exactly k of M specified numbers will be among a sample of n balls, withdrawn from the urn.

Solution. The exhaustive number of cases is $t = \binom{N}{n}$. Specification partitions the set into two subsets consisting of $N - M$ and M numbers. If k numbers are from the specified M numbers, these can be selected in ${}^MC_k$ ways. The rest of $n - k$ numbers must be selected out of $N - M$ numbers in $\binom{N-M}{n-k}$ ways.



Hypergeometric model

By the Sequential Principle of Counting, the total number of favourable cases is f

$$f = \binom{M}{k}\binom{N-M}{n-k}, \text{ whence } p = \frac{f}{t} = \binom{M}{k}\binom{N-M}{n-k}\Big/\binom{N}{n}.$$

***Example* 3.** An urn contains N balls numbered 1 through N, where the first D balls are defective and the remaining $N - D$ are non-defective. A sample of size n balls is drawn from the urn. Let A_k be the event that the sample of n balls contains exactly k defectives. Find $P(A_k)$ when the sample is drawn (i) without replacement and (ii) with replacement.

Solution. (i) Suppose numbers on the n balls are $v_1, v_2, \ldots, v_n$ and take $\Omega = [\{v_1, v_2, \ldots, v_n\}, v_j$ are numbers on n balls] so that Ω is made up of *subsets* of n. The number of subsets of size n of the N balls is $t = \binom{N}{n}$ and all such selections may be treated equi-probable. Now A_k consists of those subsets of size n which contain exactly k defective balls from the balls that are numbered 1 to D inclusive. Such a selection can be had in $\binom{D}{k}$ ways. The remaining $(n - k)$ balls are to be drawn from balls numbered $D + 1$ to N and this selection can be made in $\binom{N-D}{n-k}$ ways. By sequential counting the favourable cases for the required event is $f = \binom{D}{k}\binom{N-D}{n-k}$ ways. Hence

$$P(A_k) = \frac{f}{t} = \binom{D}{k}\binom{N-D}{n-k}\bigg/\binom{N}{n}.$$

(ii) Under replacement scheme, the total number of drawing n balls out of N balls in $(N)^n$, since at each draw, the ball can be one from 1 to N inclusive. Let v_j be the number of the ball drawn on the jth draw so that $\Omega = \{(v_1, \ldots, v_n)\}$. Now A_k consists of exactly k of the v_j's balls numbered from 1 to D inclusive; these k balls can be selected in $\binom{n}{k}$ ways to fill k positions out of n available positions. For each different $\binom{n}{k}$ selection, there are $D^k (N - D)^{n-k}$ different n-tuples (listing) by Sequential Counting. Hence

$$P(A_k) = \binom{n}{k} D^k (N-D)^{n-k}\Big/(N)^n = \binom{n}{k} q^{n-k} p^k. \qquad [p = D/N, q = 1 - p]$$

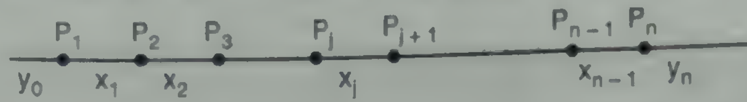
***Example* 4.** A car C is parked among N cars in a row, not at either end. On his return, the owner finds that exactly r of the N places are still occupied. What is the probability that both neighbouring places are empty ?

Solution. On his return, the owner finds that exactly r of N places are still occupied. This includes the space occupied by car C as well. This means, the remaining $(r - 1)$ cars are parked in the $(N - 1)$ slots. This provides the total choices $t = \binom{N-1}{r-1}$.

Since the neighbouring slots to car C are empty, it follows that the remaining $(r - 1)$ cars are parked among the $(N - 3)$ slots. This provides the favourable choices as $f = \binom{N-3}{r-1}$. Consequently,

$$p = \frac{f}{t} = \binom{N-3}{r-1}\bigg/\binom{N-1}{r-1} = \frac{(N-3)!(N-r)!}{(N-1)!(N-r-2)!} = \frac{(N-r)(N-r-1)}{(N-1)(N-2)}.$$

***Example* 5.** Out of m persons sitting in a row, n are chosen at random. Find the probability that no two of the chosen persons were seated together.



Solution. Suppose the n persons chosen are $P_1, P_2, \ldots, P_n$. The number of such selections of n persons out of m is $N = \binom{m}{n}$.

Let y_0 be the number of persons to the left of P_1 and y_n be the number of persons to the right of P_n (the last person). Let x_j be the number of persons between P_j and P_{j+1}. Then,

$y_0 \geq 0,\ y_n \geq 0$, but $x_j \geq 1,\ 1 = j \leq n - 1$.

$\therefore\ y_0 + (x_1 + x_2 + \ldots + x_{n-1}) \neq y_n = m - n.$ (Number of unchosen persons).

Let $x_0 = y_0 + 1,\ x_n = y_n + 1$, so that $x_j > 0,\ j = 0, 1, \ldots, n$.

The preceding equation now reduces to

$$x_0 + x_1 + \ldots + x_{n-1} + x_n = m - n + 2, \qquad x_j > 0 \text{ (strictly)}.$$

The number of solutions of this equation, in positive integers, is

$$f = \binom{(m-n+2)-1}{(n+1)-1} = \binom{m-n+1}{n}$$

Thus $p = \dfrac{f}{N} = \dbinom{m-n+1}{n} \Big/ \dbinom{m}{n}$.

***Example* 6.** Out of N tickets consecutively numbered, 3 are drawn at random. Find the chance that the numbers on them are in arithmetical progression when (i) $N = 2n$, (ii) $N = 2n + 1$.

Solution. In any choice of 3 members in A.P., say, $a - d, a, a + d$, the sum of extremes is always even $[(a - d) + (a + d) = 2a]$, being *twice* of the middle term. Two such numbers, say, x and y, whose sum is even must be drawn either from E (even-numbered set) or from O (odd-numbered set).

(i) $N = 2n,\ O = \{1, 3, 5, \ldots 2n - 1\},\ E = \{2, 4, \ldots, 2n\}$,

As $x \in O, y \in O$ or $x \in E, y \in E$

$$\therefore \qquad f = \binom{n}{2} + \binom{n}{2} = 2\binom{n}{2} = n(n-1).$$

The exhaustive number of cases $= \dbinom{2n}{3} = \dfrac{2n!}{3!(2n-3)!} = \dfrac{2n(2n-1)(n-1)}{3}$

$$\therefore \qquad p = \frac{3n(n-1)}{2n(n-1)(2n-1)} = \frac{3}{2(2n-1)}.$$

(ii) $N = 2n + 1$, $O = \{1, 3, 5, ..., 2n + 1\}$, $E = \{2, 4, 6, ..., 2n\}$

As $x \in O, y \in O$ or $x \in E, y \in E$,

$$\therefore \quad f = \binom{n+1}{2} + \binom{n}{2} = n^2.$$

The exhaustive number of cases $= \binom{2n+1}{3} = \dfrac{(2n+1)!}{3!(2n-2)!} = \dfrac{n(4n^2-1)}{3}$

$$\therefore \quad p = \frac{3n^2}{n(4n^2-1)} = \frac{3n}{(4n^2-1)}.$$

***Example* 7.** Compute the probability of each of the following 5-card poker hands :

(a) Royal Flush	$\{10, J, Q, K, A$: all of the same suit$\}$
(b) Straight Flush	(Five cards of the same suit in a sequence)
(c) Four of a kind	(Face values of the form x, x, x, x, y ; $y \neq x$)
(d) Full house	(Face values of the form x, x, x, y, y ; $y \neq x$)
(e) Straight	(Five cards in a sequence regardless of suit)
(f) Flush	(Five cards of the same suit, but not in a sequence)
(g) Two pairs	(Any two distinct face values occurring exactly twice).

Note. *Spade Flush* (Five spades), *Ace-high spade flush* (five spades with the ace). *Ace-high fall house* (three aces with another pair).

Solution. A poker hand is a 5-subset of the set of 52 cards in the full deck, and hence there are $N = \binom{52}{5}$ different poker hands, since order is of no consequence.

(a) There is one royal flush in each of the four suits, hence probability $= 4/N$.

(b) There are ten sequences of 5 cards, starting with the sequence 12345 and ending with the sequence 10 *JQKA* in any particular suit, say Diamonds. Hence the total number of 5-card sequences is 4×10, whence $p = 40/N$.

N.B. Some people consider straight flush devoid of royal flush. For such a case $p = (40 - 4)/N = 36/N$.

(c) Here $\{x, x, x, x, ; y\}$, e.g. (5, 5, 5, 5, 8) and it involves all the four suits. We choose the face value x to appear four times (thirteen choices : $A, K, Q, ..., 2$) and for the 5th card, there are 48 choices. Hence $f = 13 \times 48 = 624$. Thus, by Laplace definition, $p = 624/N$, or

$$P(\text{Four of a kind}) = \binom{4}{4}\binom{4}{1} \times 13 \times 12\left(\frac{1}{N}\right) = \frac{624}{N}.$$

(d) $P\{\text{Full house}\} = \binom{4}{3}\binom{4}{2} \times 13 \times 12\left(\dfrac{1}{N}\right) = \dfrac{3744}{N}.$

Here there are 13 choices for x (one value) and 12 choices for y (second value).

(e) The sequence of face value may be chosen in 10 ways. The suit of each card in a straight may be chosen in 4 ways. Thus, the suits of 5 cards may be chosen in $(4)^5$ ways and the probability of a Straight is, thus $p = 10 \times 4^5/N$.

with replacement — means the same item can be chosen more

(f) Five cards from the diamond-suit can be chosen in $\binom{13}{5}$ ways and since there are four suits, the total number of choices from all the four suits is $4\times\binom{13}{5}$. Hence

$$p = 4\times\binom{13}{5}\Big/N = 0.002.$$

This is the flush of any kind. Now to eliminate the possible sequences in these 5 cards, we need to have, using (b), the favourable cases

$$4\times\binom{13}{5} - 40 = 5148 - 40 = 5108 \Rightarrow p = 5108/N.$$

(g) Two pairs : $(x, x; y, y, z)$ e.g. (2, 2 ; 9, 9 ; 5). The two face values x, y can be chosen out of 13 designations $\{2, 3, J, Q, K, A\}$ in $\binom{13}{2}$ ways. For each of the two designations, we can choose the two suits in $\binom{4}{2}$ ways. The remaining 5th card z (e.g. 5) has to be chosen from 44 cards [52 – [2, 2, 2, 2, 9, 9, 9, 9] in $\binom{44}{1}$ ways. Thus,

$$P\{\text{Exactly two pairs}\} = \binom{13}{2}\binom{4}{2}^2\binom{44}{1}\Big/N.$$

Problems with Solutions Provided at the End of the Text

1*. At a bridge (game) hand North and South have n cards of one suit (trumps) S. Find the probability that one of their opponents (East or West) will have exactly k cards of that suit and the other will have the rest.

2*. From a set of 17 cards, numbered 1, 2, ..., 17 ; one is drawn at random. Show that the chance that the number is divisible by 3 or 7 is 7/17.

3*. What is the chance that (i) a leap year, (ii) a non-leap year; selected at random will contain 53 Sundays ?

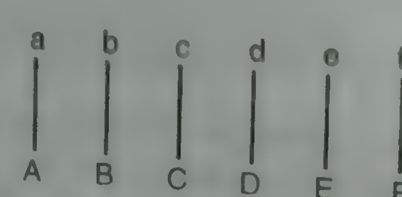
4*. A bag contains 50 tickets numbered 1, 2, ..., 50. Five tickets are drawn at random and arranged in ascending oder of magnitude : $x_1 < x_2 < \ldots < x_{50}$. What is the probability that $x_3 = 30$?

5*. What is the probability of getting 9 cards of the same suit in one hand at a game of bridge ?

6*. An urn contains 4 white, 4 red and 4 blue balls. Three balls are drawn. Find the chance that all three colours are represented when (i) sampling with replacement, and (ii) without replacement.

7*. If three small squares are chosen at random on a chess-board, find the chance that they are in a diagonal line.

8*. A girl holds six strings in her hand with the ends protruding above and below and her friend ties together the six upper ends in pairs and then ties together the six lower ends in pairs. What is the probability that at least one ring will be formed ?



without replacement — means the same item cannot be selected more than once

... order does not matter. Permutation ...

9*. Given a group of 4 people, construct an event space from which it is possible to find the probability that :

(a) Exactly two have the same birth month (event E_2).

(b) At least two have the same birth month (event L_2).

(c) No two have the same birth month (event D : different).

Assume that the probability that a person have a particular birth month is 1/12.

10*. Suppose n balls are distributed at random into N boxes. What is the probability that :

(a) No box has more than one ball ?

(b) A specified ball is in a specified box ?

11*. Out of $3n$ consecutive numbers, 3 are selected at random. Find the chance that their sum is divisible by 3.

12*. A bag contains 8 black, 3 red and 9 white balls. If 3 balls are drawn at random, find the probability that (a) all are black, (b) 2 are black and 1 is white, (c) 1 is of each colour, (d) the balls are drawn in the order black, red and white.

13*. A and B stand in a ring with 10 other persons. If the arrangement of the 12 persons is at random, find the chance that there are exactly 3 persons between A and B.

14*. A and B stand in a ring with n other persons. If the arrangement of the $(n + 2)$ persons is at random, find the chance that there are exactly k persons between A and B.

15*. Four girls a, b, c, d line up in random order in front of D.T.C. Pass Section. What is the chance that a precedes b and c ?

16*. An urn contains n red and m black balls. All the balls are withdrawn one at a time. Find the chance that no two consecutive black balls are witdrawn, $n \geq m - 1$.

17*. An urn contains w white and b black balls. The balls are drawn randomly one by one until those of the *same colour* are left. What is the probability that they are white ?

18*. From a pack of 52 cards, an even number of cards is drawn. Show that the probability of half of these cards being red is $\dfrac{[52!/(26!)^2]-1}{(2^{51}-1)}$.

19*. A and B stand in a row with n other persons. If the arrangement of the $(n + 2)$ persons is at random, find the chance that there are exactly k persons between A and B.

1-42. *Random Variables or Variates*

Very old fashioned but still very useful and most intuitive definition describes a *random variable* (r.v.) as a variable that takes on its value by chance. This older definition just stated serves quite adequately, in almost in all instances of probability modelling. In fact, this older definition has served over a century of meaningful progress in probability theory. [Modern approach appears in Chapter 3 onwards].

Conventionally, we use capital letters such as X, Y, Z to denote the r.vs. and lower case letters such as x, y, z for real numbers that such variates assume as their values. The probability that the event $\{X \leq x\}$ occurs, is stated as

$$F(x) = P\{X \le x\}, \quad -\infty < x < \infty \qquad ...(1)$$

and is called the *cumulative distribution function* of the r.v. X, abbreviated c.d.f. (or even d.f.). Note that $F(x)$ contains all the information available about the r.v. X. For instance

$$P\{X > x\} = 1 - F(x), \; P\{a < X \le b\} = F(b) - F(a)$$

$$P\{X = a\} = F(a) - \lim_{\varepsilon \downarrow 0} F(a - \varepsilon) = F(a) - F(a-).$$

If there is a finite or denumerable set of distinct values x_1, x_2, such that $P(X = x_i) = p_i > 0$, for $i = 1, 2, ..., \sum p_i = 1$, then X is called *discrete* r.v. and p_i is called *its probability mass function* (p.m.f). If $p\{X = x\} = 0$, $\forall\ x \in R$, then r.v. X is called *continuous*. If there is a non-negative function $f(x)$, defined for $-\infty < x < \infty$ such that

$$P\{a < X \le b\} = \int_a^b f(x)\,dx \quad -\infty < a < b < \infty$$

then $f(x)$ is called the *probability density function* (p.d.f) for the r.v. X.

There is a galaxy of standard p.m.fs. and p.d.fs., which describe various phenomena and are described in the succeeding chapters.

Bernoulli Random Variable. A variate X such that $P\{X = 1\} = p$ and $P\{X = 0\} = 1 - p$ is called *Bernoulli variate*. It is frequently used as **indicator** of an event A, written $\mathbb{I}(A)$ or $\mathbb{I}_A$ and defined by

$$\mathbb{I}_A = \begin{cases} 1, & \text{if } A \text{ occurs, probability } R \\ 0, & \text{otherwise} \end{cases}$$

1-43. *Expected Values*

If X is a r.v. and g is a function, then $g(X)$ is also a r.v. If X is a discrete r.v. with possible values $x_1, x_2, ..., x_i, ...$, then the expected value of $g(X)$ is denoted by $E\{g(X)\}$ and is defined by

$$E\{g(X)\} = \sum_i p(x_i) \,.\, g(x_i) \qquad ...(1)$$

provided the sum in (1) converges *absolutely*. If X is continuous with p.m.f. : $f(x)$, then $E\{g(X)\}$ is defined by

$$E\{g(X)\} = \int_{-\infty}^{\infty} f(x).g(x)\,dx \qquad ...(2)$$

provided the integral is *absolutely* convergent.

The special cases : $g(X) = X, X^2, ..., X^n$ frequently occur. Thus

$E(X) = \sum p(x_i)\, x_i, \;......,\; E(X^n) = \sum p(x_i)\, x_i^n,\; i = 1, 2, ...$ $\qquad$ (X : discrete)

$E(X) = \int_{-\infty}^{\infty} f(x)\, x\, dx, \;......,\; E(X^n) = \int_{-\infty}^{\infty} f(x)\, x^n\, dx,$ $\qquad$ (X : continuous)

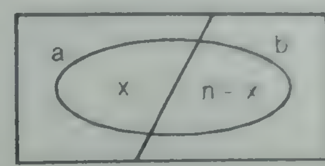
The quantity $E(X)$ is usually denoted by μ and is called expected value of X or mean value of X or simply mean (of X). This occurs most frequently in probability theory. The most valuable property of expectation is the linear property of E (written Lin E) and is given by

$$E\,[ag(X) + bh(X) + c] \;= aE[g(X)] + bE[h(X)] + c$$

where a, b, c are constants and g, h are arbitrary functions.

Comments. A specific p.m.f., called *hypergeometric distribution*, is defined by

$$P\{X = n\} = \binom{a}{x}\binom{b}{n-x}\bigg/\binom{a+b}{n},\; x = 0, 1, 2, \ldots, n$$



We often write it as $X \sim$ HG $(a + b, a, n)$

It is proved later on, that

$$E(X) \;= np, \text{ where } p = a/(a + b).$$

We use this result in Chapter 2 to offer shorter evaluations.

On use of Indicators

The Bernoulli r.v. X with parameter p, identified as indicator of an event A yields

$$E(I_A) \;= 1\,.\,p + 0\,.\,(1 - p) = p = P(A)$$

Thus, $\quad P(A) \;= E(I_A)$

This simple expedient of using Indicators frequently reduces formidable calculations into trivial evaluations.

Illustration. Show that $\sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n} a_i\,a_j\,p_{ij} \geq 0$. $\;[p_{ij} = P(A_iA_j)]\,a_k \in R.$

To prove it, we note that

$$0 \leq \left(\sum_{i=1}^{n} a_i\,I_{A_i}\right)^2 \;= \left(\sum_{i=1}^{n} a_i\,I_{A_i}\right)\left(\sum_{j=1}^{n} a_j\,I_{A_j}\right)$$

$$= \sum_i\sum_j a_i\,a_j\,I_{A_i}\,I_{A_j} = \sum_i\sum_j a_{ij}\,I_{A_iA_j}$$

Taking expectation of both sides, using $E(I_{A_iA_j}) = P(A_iA_j)$ yields the stated result trivially.

A direct attack to evaluate is indeed very difficult.

An excellent use of indicators is shown in getting Poincare's inequalities [See §1-53]

1-44. *Two-Dice Experiments*

Two fair dice are tossed simultaneously, their outcomes $\omega = (x, y)$ are recorded. Let S, D, T, M, m denote sum, difference, product, maximum and minimum of the eyes shown. Record the sample space and their joint probability distributions. The probability of each of the 36 outcomes is $p = 1/36$. The individual probability mass functions and joint mass functions follow this table :

$\omega = (x, y)$	$S = x + y$	$D = x - y$	$\lvert D \rvert$	$T = xy$	Max ω	Min ω
(1, 1)	2	0	0	1	1	1
(1, 2)	3	−1	1	2	2	1
(1, 3)	4	−2	2	3	3	1
(1, 4)	5	−3	3	4	4	1
(1, 5)	6	−4	4	5	5	1
(1, 6)	7	−5	5	6	6	1
(2, 1)	3	1	1	2	2	1
(2, 2)	4	0	0	4	2	2
(2, 3)	5	−1	1	6	3	2
(2, 4)	6	−2	2	8	4	2
(2, 5)	7	−3	3	10	5	2
(2, 6)	8	−4	4	12	6	2
(3, 1)	4	2	2	3	3	1
(3, 2)	5	1	1	6	3	2
(3, 3)	6	0	0	9	3	3
(3, 4)	7	−1	1	12	4	3
(3, 5)	8	−2	2	15	5	3
(3, 6)	9	−3	3	18	6	3
(4, 1)	5	3	3	4	4	1
(4, 2)	6	2	2	8	4	2
(4, 3)	7	1	1	12	4	3
(4, 4)	8	0	0	16	4	4
(4, 5)	9	−1	1	20	5	4
(4, 6)	10	−2	2	24	6	4
(5, 1)	6	4	4	5	5	1
(5, 2)	7	3	3	10	5	2
(5, 3)	8	2	2	15	5	3
(5, 4)	9	1	1	20	5	4
(5, 5)	10	0	0	25	5	5
(5, 6)	11	−1	1	30	6	5
(6, 1)	7	5	5	6	6	1
(6, 2)	8	4	4	12	6	2
(6, 3)	9	3	3	18	6	3
(6, 4)	10	2	2	24	6	4
(6, 5)	11	1	1	30	6	5
(6, 6)	12	0	0	36	6	6

We now condense this data in various forms and prepare tables for ready-reference for future use.

1. ***Distribution of Sum.*** Let S_k denote $x + y = k$. Then the above table yields : $P(S_2) = p$, $P(S_3) = 2p$, $P(S_4) = 3p$, $P(S_5) = 4p$, $P(S_6) = 5p$, $P(S_7) = 6p$, $P(S_8) = 5p$, $P(S_9) = 4p$, $P(S_{10}) = 3p$, $P(S_{11}) = 2p$, $P(S_{12}) = p$.

Summary. $\boxed{P\{S_R\} = 6 - |7 - k|/36,\ k = 2, 3, \ldots, 11, 12}$

2. ***Distribution of Maximum.*** Let M_k denote that max. $(x, y) = k$. Then table yields $P(M_1) = p$, $P(M_2) = 3p$, $P(M_3) = 5p$, $P(M_4) = 7p$, $P(M_5) = 9p$, $P(M_6) = 11p$.

Distribution of the Maximum

$x_i\downarrow$ \ $M_j\rightarrow$	1	2	3	4	5	6	$P(x_i)$
1	p	p	p	p	p	p	$6p$
2	–	$2p$	p	p	p	p	$6p$
3	–	–	$3p$	p	p	p	$6p$
4	–	–	–	$4p$	p	p	$6p$
5	–	–	–	–	$5p$	p	$6p$
6	–	–	–	–	–	$6p$	$6p$
$P(M_j)$	p	$3p$	$5p$	$7p$	$9p$	$11p$	$36p$

Distribution of the Minimum

$x_i\downarrow$ \ $m_j\rightarrow$	1	2	3	4	5	6	$P(x_i)$
1	$6p$	–	–	–	–	–	$6p$
2	p	$5p$	–	–	–	–	$6p$
3	p	p	$4p$	–	–	–	$6p$
4	p	p	p	$3p$	–	–	$6p$
5	p	p	p	p	$2p$	–	$6p$
6	p	p	p	p	p	p	$6p$
$P(m_j)$	$11p$	$9p$	$7p$	$5p$	$3p$	p	$36p$

Distribution of Max. and Min.

$m_j\downarrow$ \ $M_i\rightarrow$	1	2	3	4	5	6	$P(m_j)$
1	p	$2p$	$2p$	$2p$	$2p$	$2p$	$11p$
2	–	p	$2p$	$2p$	$2p$	$2p$	$9p$
3	–	–	p	$2p$	$2p$	$2p$	$7p$
4	–	–	–	p	$2p$	$2p$	$5p$
5	–	–	–	–	p	$2p$	$3p$
6	–	–	–	–	–	p	p
$P(M_j)$	p	$3p$	$5p$	$7p$	$9p$	$11p$	$36p$

Distribution of sum and abs. difference

$\lvert D_i \rvert\downarrow$ \ $S_j\rightarrow$	2	3	4	5	6	7	8	9	10	11	12	$P(D_i)$
0	p	–	p	–	p	–	p	–	p	–	p	$6p$
1	–	$2p$	–	$2p$	–	$2p$	–	$2p$	–	$2p$	–	$10p$
2	–	–	$2p$	–	$2p$	–	$2p$	–	$2p$	–	–	$8p$
3	–	–	–	$2p$	–	$2p$	–	$2p$	–	–	–	$6p$
4	–	–	–	–	$2p$	–	$2p$	–	–	–	–	$4p$
5	–	–	–	–	–	$2p$	–	–	–	–	–	$2p$
$P(S_j)$	p	$2p$	$3p$	$4p$	$5p$	$6p$	$5p$	$4p$	$3p$	$2p$	p	$36p$

Distribution of Sum and Product

$S\downarrow$ \ $T\rightarrow$	1	2	3	4	5	6	8	9	10	12	15	16	18	20	24	25	30	36	$P(S_i)$
2	p	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	p
3	–	$2p$	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	$2p$
4	–	–	$2p$	p	–	–	–	–	–	–	–	–	–	–	–	–	–	–	$3p$
5	–	–	–	$2p$	–	$2p$	–	–	–	–	–	–	–	–	–	–	–	–	$4p$
6	–	–	–	–	$2p$	–	$2p$	p	–	–	–	–	–	–	–	–	–	–	$5p$
7	–	–	–	–	–	$2p$	–	–	$2p$	$2p$	–	–	–	–	–	–	–	–	$6p$
8	–	–	–	–	–	–	–	–	–	$2p$	$2p$	p	–	–	–	–	–	–	$5p$
9	–	–	–	–	–	–	–	–	–	–	–	–	$2p$	$2p$	–	–	–	–	$4p$
10	–	–	–	–	–	–	–	–	–	–	–	–	–	–	$2p$	p	–	–	$3p$
11	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	$2p$	–	$2p$
12	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	p	p
$P(T_j)$	p	$2p$	$2p$	$3p$	$2p$	$4p$	$2p$	p	$2p$	$4p$	$2p$	p	$2p$	$2p$	$2p$	p	$2p$	p	$36p$

Note. These tables can be adopted for tetrahedral dice by considering $1 \le x, y \le 4$.

Exercise 1(a)

1. An urn contains 30 identical balls numbered from 1 to 30. If a ball is selected at random, show that the probability that the number on the selected ball is (a) an even number is 1/2, (b) an odd number is 1/2, (c) a number divisible by 5 is 1/5.
2. Two dice are thrown. Find the probability that the sum of scores on the dice is (a) 9, (b) greater than 8, (c) either less than 8 or odd.
   Verify that chance of throwing more than 7 equals that of throwing less than 7, each being 5/12.
3. A bag contains 5 red, 6 black and 3 white balls. Three balls are drawn at random. Show that the probability that there is one ball of each colour is 45/182.
4. (a) From a pack of 52 cards, three are drawn at random. Show that the chance that they are a king, a queen and a knave is 16/5525.
   (b) The 52 cards of a deck are placed successively one after the other and from left to right. Find the chance that the 13th Heart will appear before the 13th Spade.
5. Two cards are drawn from an ordinary pack. Show that the probability that
   (a) Both are aces is 1/221, (b) One is a king and the other a queen is 8/663, (c) One is a spade and the other a heart is 13/102.
6. A number is chosen from each of the two sets A and B.
   $$A = \{1, 2, 3, 4, 5, 6, 7, 8, 9\}\ ;\quad B = \{4, 5, 6, 7, 8, 9\}.$$
   If p_1 is the probability that the sum of two numbers be 10 and p_2 the probability that their sum is 8, find $p_1 + p_2$.
7. Six faces of a die are marked with the numbers $1, -1, 0, 2, -2, 3$. If the die is thrown thrice, what is the probability that the sum of numbers thrown will be equal to 6.
8. Two different digits are chosen at random from the set $\{1, 2, 3, ..., 8\}$. Show that the probability that the sum of the digits equals 5 is the same as the probability that their sum exceeds 13, each being 1/14. Also show that the chance of both digits exceeding 5 is 3/28.
9. (a) Two dice are thrown. It is given that one of the dice shows 5 points. Show that the probability that the sum of points on the two dice being 9 is 1/6.
   (b) Two dice are cast. Person A wins if the sum of numbers showing up is ≤ 6 and one of the dice shows 4. Person B wins if the sum is ≥ 5 and one of the dice shows 4. Prove that $P(A \text{ wins}) = 4/36$, $P(B \text{ wins}) = 11/36$ and $P(\text{both } A \text{ and } B \text{ win}) = 4/36$.
10. (a) A number of two digits is chosen at random from numerals $1, 2, ..., 9$. Show that the probability that the number so chosen is even and less than 60 is 1/4.
    (b) Urns A and B contains n chips numbered 1 to n; one chip is drawn from A as well as from B. Show that $P\{\text{chip drawn from } B \text{ bears a number smaller than that from } A\} = (n-1)/2n$.
11. From 100 tickets numbered 1, 2, ..., 100, four are drawn at random. What is the probability that three of them will bear numbers 1 to 20 and the fourth will bear any number from 21 to 100 ?
12. The first 12 letters of English alphabet are written down at random. Show that the probability that there are exactly four letters between A and B is 7/66.
13. An urn contains M balls numbered 1 to M. A sample of size n is drawn without replacement, and balls are arranged in increasing order of their numbers $x_1 < x_2 < ... < x_n$. Let K be a number from 1 to M and k a number from 1 to n. Show that
    $$P\{X_k = K\} = \binom{K-1}{k-1}\binom{M-K}{n-k} \Big/ \binom{M}{n}.$$

14. Four persons are chosen at random from a group containing 3 men, 2 women and 4 children. Show that the chance that exactly two of the them will be children is 10/21.

15. In a game of whist (where 4 players play), prove that the chance that the 4 kings are held by a specified player is 11/4165.

16. In a Bridge-party, North has four aces. Find the probability that he gets a king of spade.

17. Show that the probability that at a Bridge game, (i) a hand holds 4 aces is 11/4165, (ii) a hand holds exactly six spades is $39!\,(13\,!)^2/52\,!\,7\,!\,6\,!$.

18. Let n biscuits be distributed among N beggars. Find the chance that a particular beggar receives $r\ (< n)$ biscuits ?

19. m men and w women seat themselves at random on $m + w$ seats arranged in (i) a row, (ii) circle. Find the probability that all the women will be adjacent. $[p_1 = (m+1)/{}^{(m+w)}C_w]$.

20. Each coefficient in the equation, $ax^2 + bx + c = 0$, is determined by throwing an ordinary die. Obtain the probability that the equation has (i) real roots, (ii) equal roots, (iii) rational roots.

21. $2n$ people are queued up at a theatre box office; n of them have only 5-rupee bill and the remaining n only one 10-rupee bill. There is no cash in the box office when it opens and each customer in turn is going to purchase a single 5-rupee ticket. Show that the probability that no customer will have to wait for change is $1/(1 + n)$.

Re-arrangements of Words :

22. Letters are drawn one at a time from a box containing the letters : A, H, M, O, S, T. Show that the probability that the letters in the order drawn spell the word 'THOMAS' is 1/720.

23. Letters of the word LATENT are written on six cards. The cards are shuffled and then drawn one at a time at random. Show that the probability that the word 'TALENT' is formed is 1/360.

24. A letter is taken out at random out of "ASSISTANT" and a letter is taken out of 'STATISTICS'. Show that the chance that they are the same letter is 19/90.

25. If the letters of the word "REGULATIONS" be arranged at random, show that the probability that there will be exactly 4 letters between R and E is 6/55.

26. Show that the probability that in a random arrangement of letters of the word "UNIVERSITY", the two I's do not come together is 4/5.

27. In a random arrangement of the letters of the word COMMERCE, find the probability that all the vowels come together.

1-45. *Shortcomings of Classical Approach to Probability*

In spite of its elegance and utility, the classical definition suffers from some serious defects. Some of the drawbacks are hinted below :

1. The definition of classical probability is circular, in the sense that 'equilikely' and 'equiprobable' are synonymous terms which form the definition unacceptable. Because of this difficulty in deciding about the equilikely alternatives, serious errors are committed. In this context, we give an example, which is now known as *D'Alembert's Paradox* : Let two coins be tossed. D'Alembert argued that there are three possible cases, viz. : (i) Both heads, (ii) Both tails, (iii) One head and one tail, and he thus concluded that the probability of a head and a tail is 1/3, which is obviously not true. In fact, the actual probability is 2/4 = 1/2.

We may conclude that the wrong decision about the equilikely cases can lead to errors, committed even by great Mathematicians.

2. In many problems, the possible number of outcomes is infinite. For instance, the intuitive answer, on the basis of classical definition, to the probability that a positive integer chosen at random is *even* is 1/2. The plausible argument, that out of first $2n$ positive integers, n are *even*, so $p = n/2n = 1/2$ (fixed), even as $n \to \infty$ (so as to cover all positive integers) is not an easy matter to make it acceptable. For one thing, it depends on natural ordering of positive integers and so a different ordering would yield a different result. For example, for the ordering : 1, 3, 2, 5, 7, 4, 9, 11, 6 ; $p = 1/3$. In fact, just by choosing some ordering, we could obtain any probability that we desired. It follows that, the classical definition does require modification.

3. Since classical definition is based on equiprobable events, it naturally fails to solve problems related to events which are not equiprobable. For example, for a loaded coin which is biased in favour of head, the classical definition fails to give correct answer, since the outcomes are not equally-likely to occur.

4. The classical definition offers very limited uses. It fails to respond to querries like :

What is the probability that an Indian shall live upto the age of 80 years ? What is the chance that a bulb will burn 100 hours or more ? What is the chance that two cars going through the same intersection will collide ?

These questions cannot be dealt with by using classical definition because 'symmetry' and 'equilikelyness' cannot be guessed as in games of chance.

Thus, we need to extend the classical definition in such a way that the new definition is able to overcome the various defects embodied in the classical-definition approach. To remedy the above deficiencies and the like, we introduce the *relative frequency* approach in the following section.

1-46. *Relative Frequency Approach to Probability*

In the beginning of the 20th Century, Richard Von Mises introduced the relative frequency approach by such considerations as under :

Let an experiment under consideration be repeated N times and suppose that an event A occurs $n(A)$ times. The ratio $n(A)/N$ is termed the *relative frequency* of the event A. As the number of trials N increases, the relative frequency of the event A approaches (stabilizes) the number $P(A)$, called the *probability of event* A; i.e. $P(A) = \lim n(A)/N$. $(N \to \infty)$.

For example, to predict whether the next baby born in a certain locality will be male or female, which is individually an uncertain event, we search for the long-run records. If the previous records show that 52% of the births are male, it is safe to postulate that the probability of a male birth in this locality is p and approximate p as 0.52.

We may note that the laws of addition and multiplication for probability can be proved by *relative frequency approach* as well.

Although the Frequentist Approach overcomes some of the deficiencies of Laplace classical definition, it has some demerits of its own.

In the first place, it is more in the way of a description rather than of a formal mathematical character. In physical experiments, the number N might be large but it is always finite ; therefore $n(A)/N$ cannot be equated even approximately to a limit. In general, the relative frequency of an event tends to its probability but in its own manner not for certain, but *almost for certain* with a very high probability.

The two types of probabilities (aprori and posteriori) have one important thing in common. They both require conceptual experiment in which the various outcomes can occur under somewhat uniform conditions. But we can have the situations where we cannot fit in the framework of repeated trials. For example, what is the probability that my wife loves me ? Or what is the probability that Third World War will start before Ist Jan. 2000 ? These types of problems are certainly a legitimate part of general probability theory and are titled "Subjective Probabilities". We shall not discuss the subjective probabilities in this work, but the axioms of probability from which we develop the theory are rich enough, to include a priori probability, a posteriori probability and subjective probabilities as well.

1-50. *Axiomatic Treatment of Probability Theory*

***Definition* 1.** Let Ω be any non-empty set. A non-empty class $\mathcal{F}$ of subsets of Ω called a σ-*field* (or σ-*algebra*) if it satisfies the following conditions :

(i) $\Omega \in \mathcal{F}$ (ii) $A \in \mathcal{F} \Rightarrow \overline{A} \in \mathcal{F}$ (iii) $A_n \in \mathcal{F}$, $n = 1, 2, 3, \ldots$ gives $\bigcup_{n=1} A_n \in \mathcal{F}$.

***Example* 1.** Let Ω be any non-empty set and $\mathcal{F}$ denote the class of all subsets of Ω. Then $\mathcal{F}$ is a σ-field.

***Example* 2.** Let Ω be any non-empty set. Then $\mathcal{F} = \{\varnothing, \Omega\}$ is the smallest σ-field associated with Ω.

***Example* 3.** Let Ω be any non-empty set. Let $A \neq \phi$ be any subset of Ω. Then $\{\varnothing, A, \overline{A}, \Omega\}$ is a σ-field.

***Definition* 2.** Let Ω denote the sample space of an unidentified random experiment Let $\mathcal{F}$ be a σ-field associated with Ω. The elements of Ω are called *sample points* and the elements of $\mathcal{F}$ are called *events.*

***Definition* 3.** If Ω is finite or it contains a countable number of elements then the sample space Ω is called a *discrete sample space*. If Ω is uncountable, then it is called an *uncountable sample space*. In particular, if Ω is rectangle (or interval) of the Euclidean space R^n, $n \geq 1$, then Ω is called a *continuous sample space*. The pair $(\Omega, \mathcal{F})$ is termed *measurable space*.

Comments. In general, the σ-field associated with Ω need not contain all the subsets of Ω. For example, if $\Omega = \{1, 2, 3, 4, 5, 6\}$, Die-sample space, then $\mathcal{F} = \varnothing, \Omega, \{1, 2, 3\}, \{4, 5, 6\}$ is a σ-field which does not contain all the subsets of Ω, e.g. $\{1, 2\} \notin \mathcal{F}$.

In the present chapter, we shall concentrate on *discrete sample spaces* and σ-field will be class of all subset of Ω, the underlying sample space.

Example 4. Toss a coin, then $\Omega = \{H, T\}$, (H is head and T is tail). Let $\mathcal{F}$ be the σ-field consisting of all subsets of Ω. Then $\mathcal{F}$ contains $2^2 = 4$ events including 2 elementary events.

Example 5. Roll a die, then $\Omega = \{1, 2, 3, 4, 5, 6\}$. Let $\mathcal{F}$ be the σ-field consisting of all subsets of Ω. Then $\mathcal{F}$ consists of $2^6 = 64$ elements (events) including 6 elementary events.

Example 6. Toss a coin till the first head appears ; then

$$\Omega = [\{H\}, \{T, H\}, \{T, T, H\}, ...]$$

Let $\mathcal{F}$ be the σ-field of all subsets of Ω. Another way to describe Ω is that we look at the number of tosses required for the first head. Clearly, then $\Omega = \{1, 2, 3, 4, ...\}$, the set of positive integers. We then let $\mathcal{F}$ be the set of all subsets of positive integers.

Exercise 1(b)

1. (i) Give classical and statistical definitions of probability and their limitations. State axioms of probability due to Kolmogorov.

(ii) Define σ-field, minimal σ-field and monotone field. Exemplify : A field containing infinite number of points, may not be a σ-field.

2. A, B, C are three arbitrary events. Express the following events in appropriate symbols :

(a) Only A occurs (b) All three events occur (c) Both A and B, but not C occur

(d) At least one occurs (e) At least two occur (f) Two and no more occur

(g) One and no more occurs (h) None occurs

(i) Every point of A is contained in B (j) Event B, but not A occurs

(k) A, B, C are mutually exclusive events.

[Watch Fig. § 1.54, Rider : $A\bar{B}\bar{C}$, ABC, $AB\bar{C}$, $L_1 = A \cup B \cup C$, $L_2 = \bar{A}BC \cup A\bar{B}C \cup AB\bar{C} \cup ABC$

$E_2 = \bar{A}BC \uplus A\bar{B}C \uplus AB\bar{C}$, $E_1 = A\bar{B}\bar{C} \uplus \bar{A}B\bar{C} \cup \bar{A}\bar{B}C$, $E_0 = \bar{A}\bar{B}\bar{C}$, $A \subseteq B$, $\bar{A}B$, $A \cap B \cap C = \varnothing$].

3. For the experiment of throwing a die two times, describe the sample space and list all the elementary events in $A \cup B$ and $A \cap B$, where A = Number one in the first throw, B = Number one in the second throw.

4. Enumerate the elementary events and give samples space when :

(i) Two coins are tossed, (ii) Two dice are thrown.

In each of the above cases, give two examples of composite events.

5. A coin is tossed 4 times. Define a sample space Ω which describes this experiment and give the subsets of Ω which correspond to the following events :

(a) More heads than tails are obtained,

(b) Tails occur on the even numbered tosses,

(c) Heads occur on the third toss.

6. A coin is tossed until it comes up with the same side twice in succession. To each possible outcomes requiring n tosses, assign the probability $p = 2^{-n}$. Describe S : the space of elementary events and evaluate $P(A)$ and $P(B)$ where A = {Trial ends at the 6th toss}, B = {Even number of tosses are needed}. **[Ans.** $S = \{tt, hh, thh, htt, thtt, hthh, ...\}$; $P(A) = 15/16$, $P(B) = 2/3$]

7. Describe sample space appropriate in each of the following cases. How many points are there in the sample spaces ?

(i) n-tosses of a coin with head or tail as outcome in each toss,

(ii) Successive tosses of a coin until a head turns up,

(iii) A survey of families with two children is conducted and the sex of the children (the older child first) is recorded,

(iv) Two successive draws : (a) with replacement, (b) without replacement, from a bag containing 4 coloured toys, out of which one white, one black and 2 red toys.

8. A psychological experiment consists of submitting a subject to two tests and assigning for each test a score of 1, 2, 3, or 4 according to his performance. Define a suitable sample space Ω for this experiment and identify the events A, B, C which denote the results :

(i) At least one score is greater than 2.

(ii) The first score is at least double the second.

(iii) The difference in scores is exactly 2.

Identify the events (i) $A \cup C$, (ii) $\overline{A \cup B}$, (iii) $A \cup \overline{C}$. Are any two of the events $\overline{A}$, B and C disjoint ?

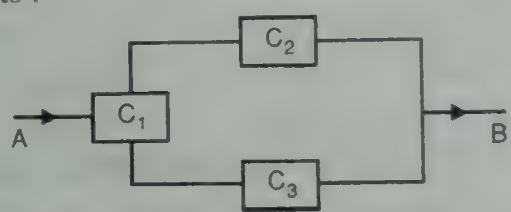
9. An assembly of electronic equipment consists of three components arranged in the series-parallel circuit (see figure).

Each component is either operative (1) or fails (0). However the entire assembly will operate iff there is unbroken path from A to B. Denote the events :

E_1 : the assembly is operative,

E_2 : C_2 has failed but the assembly is operative,

E_3 : C_3 has failed but the assembly is operative.



(i) List the elements of Ω, E_1, E_2, E_3.

(ii) How many elementary events are there ?

(iii) Which of the following pairs of events are disjoint : E_1, E_2 ; E_1, E_3; E_2, E_3 ?

1-51. *Probability Set Function*

Let Ω be a given sample space and let $\mathcal{F}$ be a σ-field associated with Ω. A probability function (or measure) P is a real-valued set function having domain $\mathcal{F}$, which assigns to each event $A \in \mathcal{F}$ a real number $P(A)$ and Codomain [0, 1] and further satisfies the following three axioms :

[$P1$] $P(A) \geq 0$, for every $A \in \mathcal{F}$. [Non-negativity]

[$P2$] $P(\Omega) = 1$, i.e. P is normed. [Normality]

[$P3$] If $A_1, A_2, \ldots, A_n \ldots$ are disjoint events then $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$.

[σ-additivity or countable Additivity]

Thus the probability function is a normed measure on the measurable space $(\Omega, \mathcal{F})$. P is also called a *Probability Measure*. The triplet $(\Omega, \mathcal{F}, P)$ is called a *Probability Space* The pair $(\mathcal{F}, P)$ is called a *Probability Field*.

Comments. When there are only a denumerable number of possible events, say, $\Omega = \{\omega_1, \omega_2, \ldots\}$ we may take $\mathcal{F}$ to be the collection of all subsets of Ω. However, if Ω is nondenumerably infinite, it may not be possible to define a probability measure on the collection of all subsets maintaining the properties [$P1$, $P2$, $P3$]. Every prescription of $\mathcal{F}$, satisfying above axioms, should satisfy (i) $\varnothing \in \mathcal{F}$, $\Omega \in \mathcal{F}$, (ii) $\overline{A} \in \mathcal{F}$ if $A \in \mathcal{F}$, (iii) $\bigcup A_j \in \mathcal{F}$, $1 \leq j < \infty$. Any collection $\mathcal{F}$ of subsets of Ω is then called σ-*Algebra*.

Example. Let P_1 and P_2 be probability functions (measures). Then the affine combination $\alpha_1 P_1(A) + \alpha_2 P_2(A) = P(A)$, say, is also a probability measure.
[Affine combination satisfies : $\alpha_1 \geq 0$, $\alpha_2 \geq 0$, $\alpha_1 + \alpha_2 = 1$. True even for $\alpha_j \geq 0$, $j = 1, 2, \ldots$ and $\sum \alpha_j = 1$, $1 \leq j < \infty$]

Solution. Given $P(A) = \alpha_1 P_1(A) + \alpha_2 P_2(A)$, setting $A = S$ yields ...(i)

$$P(S) = \alpha_1 P_1(S) + \alpha_2 P_2(S) = \alpha_1 + \alpha_2 = 1 \qquad [P_j(S) = 1, \text{ by hypothesis}] \qquad \ldots\text{(ii)}$$

Now $P(A_j) = \alpha_1 P(A_j) + \alpha_2 P(A_j)$. So

$$P\left\{\biguplus_{j=1}^{\infty} A_j\right\} = \alpha_1 P_1\left\{\biguplus_{j=1}^{\infty} A_j\right\} + \alpha_2 P_2\left\{\biguplus_{j=1}^{\infty} A_j\right\}.$$

$$= \alpha_1\left[\sum_{j=1}^{\infty} P_1(A_j)\right] + \alpha_2\left[\sum_{j=1}^{\infty} P_2(A_j)\right] \qquad [P_1 \text{ and } P_2 \text{ are probab. functions}]$$

$$= \sum_{j=1}^{\infty} \alpha_1 P_1(A_j) + \sum_{j=1}^{\infty} \alpha_2 P_2(A_j)$$

$$= \sum_{j=1}^{\infty} [\alpha_1 P_1(A_j) + \alpha_2 P_2(A_j) = \sum_{j=1}^{\infty} P(A_j) \qquad \text{[by (ii)]}$$

Thus $\quad P(\biguplus A_j) = \sum P(A_j)$, $1 \leq j < \infty$.

All three axioms : (i) Non-negativity, (ii) Normality, (iii) σ-additivity, are satisfied. It follows that P defined by (i) is also a probability measure.

Extension. Let P_j, $1 \leq j < \infty$ be probability measures. Then the affinine combination $\sum \alpha_j P(A_j)$ is a probability measure, where $\alpha_j \geq 0$, $\forall_j$ and $\sum \alpha_j = 1$, $1 \leq j < \infty$.
[The result follows as above].

1-52. *Some Consequences of Probability Axioms*

We assume an unidentified probability space in the following results :

1. $P(\varnothing) = 0$. [The probability of an impossible event is zero]

Proof. Since $\varnothing \in \mathcal{F}$, $P(\varnothing)$ is well defined. Let $A_i = \varnothing$, $i = 1, 2, \ldots$. Then A_i are certainly disjoint and $A_1 \uplus A_2 \uplus \ldots = \varnothing$. Now by σ-additivity

$$P(\varnothing) = P(A_1 \uplus A_2 \uplus \ldots) = \sum_{i=1}^{\infty} P(A_i) = \lim_{n\to\infty} \sum_{i=1}^{n} P(\varnothing) = P(\varnothing) \lim_{n\to\infty} n.$$

If $P(\varnothing) \neq 0$, this result provides $\lim n \to 1$ as $n \to \infty$, which is absurd. Hence we must have $P(\varnothing) = 0$.

Comments. $P(\varnothing) = 0$, but if $P(A) = 0$, A need not be impossible. The event A has zero probability not because it is impossible, but because it is just one of an infinite number of equally likely experimental outcomes. [Consult §1-56 Example 1]

2. ***Finite-additivity or Sub-additivity.*** If $A_1, A_2, \ldots, A_n$ are disjoint events, then

$$P(A_1 \uplus A_2 \uplus \ldots \uplus A_n) = P(A_1) + P(A_2) + \ldots + P(A_n) \qquad \ldots[P3']$$

Proof. Let $A_j = \varnothing$, for $j = n+1, n+2, \ldots$ Then, the events A_i, $i = 1, 2, \ldots$ are disjoint and

$$A_1 \uplus A_2 \uplus \ldots \uplus A_n = A_1 \uplus A_2 \uplus \ldots \uplus A_n \uplus A_{n+1} \uplus \ldots$$

$\therefore\quad P\{A_1 \uplus A_2 \uplus \ldots \uplus A_n\} = P\{A_1 \uplus A_2 \uplus \ldots \uplus A_n \uplus A_{n+1} \uplus \ldots\}$

We apply [$P3$] to the R.H.S., also use $P(\varnothing) = 0$ and obtain

$P\{A_1 \cup A_2 \cup \ldots \cup A_n\} = P(A_1) + \ldots + P(A_n) + P(\varnothing) + P(\varnothing) + \ldots = P(A_1) + P(A_2) + \ldots + P(A_n)$

This establishes the *finite additivity* of Probability function. When $n = 2$, $P(A \uplus B) = P(A) + P(B)$.

3. *Subtraction Law.* If A and B are any two two events in sample space S, then

$$\boxed{P(A\bar{B}) = P(A) - P(AB)}$$

Proof. We write $A = (A\bar{B} \uplus AB)$. [See Euler-diagram p. 3] So

$$P(A) = P\{A\bar{B} \uplus AB\} = P(A\bar{B}) + P(AB). \qquad \text{[by } P3'\text{]}$$

$$\therefore \qquad P(A\bar{B}) = P(A) - P(AB). \qquad [\text{Note}: A\bar{B} = A - B]$$

Cor. 1. Negation or Complement Rule : $\quad P(\bar{B}) = 1 - P(B)$

Take $A = S$, $S \cap \bar{B} = \bar{B}$.

Cor. 2. Monotonicity. If $B \subseteq A$, i.e. B is a sub-event of A, then $A \cap B = B$, the above result gives

$$P(A - B) = P(A) - P(B). \qquad [B \subseteq A] \qquad \ldots(\text{i})$$

Since $P(A - B) \geq 0$, [by $P1$] it follows from (i) that $P(A) \geq P(B)$. Thus

$$B \subseteq A \Rightarrow P(B) \leq P(A).$$

Cor. 3. Supremum and Infimum of Probability Values :

For any event A : $\qquad 0 \leq P(A) \leq 1. \qquad$ [Probability Scale] $\qquad \ldots(\text{ii})$

Proof. By axiom $P1$, $P(A) \geq 0$. Also $P(\bar{A}) \geq 0$ i.e. $1 - P(A) \geq 0 \Rightarrow P(A) \leq 1$.

Combining results yield (ii).

***Cor.* 4.** $P(AB) - P(A)\,P(B) = P(A')\,P(B) - P(A'B) = P(A)\,P(B') - P(AB') = P(A'B') - P(A')P(B')$

Proof. Since $\quad P(A'B) = P(B) - P(AB)$ and $P(A') = 1 - P(A)$, etc. hence

$$P(A')\,P(B) - P(A'B) = [1 - P(A)]\,P(B) - [P(B) - P(AB)] = P(AB) - P(A)\,P(B)$$

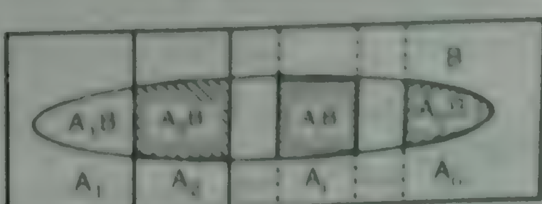
$$P(A)\,P(B') - P(AB') = P(A)\,[1 - P(B)] - [P(A) - P(AB)] = P(AB) - P(A)\,P(B)$$

$$P(A'B') - P(A')\,P(B') = [1 - P(A \cup B)]\,[1 - P(A)]\,[1 - P(B)] = P(AB) - P(A)\,P(B)$$

[by 1-52(6)]

4. *Countable Partition Property.* If $\{A_1, A_2, \ldots, A_n\}$ forms a finite partition of the sample space S, then for any event $B \subseteq S$,

$$P(B) = \sum_{i=1}^{n} P(BA_i).$$



Proof. Since $B = \biguplus_{i=1}^{n} (B \cap A_i)$, so

$$P(B) = P\left\{\biguplus_{i=1}^{n} (B \cap A_i)\right\} = \sum_{i=1}^{n} P(BA_i) \qquad \text{[by } P3'\text{]}$$

Comment. The result holds for a countable Partition $\{A_1, A_2, ..., A_n, ...\}$, since

$B = \bigcup_{i=1}^{\infty}(B \cap A_i)$ and

$$P(B) = P\left\{\biguplus_{i=1}^{\infty}(B \cap A_i)\right\} = \sum_{i=1}^{\infty} P(BA_i).$$

Cor. If S is a discrete sample sapce (finite or countable) with elementary events e_i, $k = 1, 2, 3, ...$ where e_k has probability $P(e_k)$, then for any event $E \subseteq S$, we have $P(E) = \sum_{e_i \in E} P(e_k)$.

Proof. Treat sets of S as the elementary events themselves i.e. $\{e_1, e_2, ..., e_n, ...\}$ is a partition of S.

Obviously, $e_k \cap \mathrm{E} = \begin{cases} e_k & \text{if } e_k \in \mathrm{E} \\ \varnothing & \text{if } e_k \notin \mathrm{E}. \end{cases}$ Thus, $P(\mathrm{E}) = \sum_{e_i \in S} P(\mathrm{E} \cap e_i) = \sum_{e_i \in \mathrm{E}} P(e_i)$.

5. *Union of Two Events* : $\underline{P(A \cup B) = P(A) + P(B) - P(AB).}$

Proof. Write $(A \cup B) = A \uplus \bar{A}B$ and apply finite-additivity [$P3'$] to get

$$P(A \cup B) = P\{A \uplus \bar{A}B\} = P(A) + P(\bar{A}B) = P(A) + P(B) - P(AB), \qquad \text{[by 4]}$$

6. *Union of Three Events* :

$$\underline{P(A \cup B \cup C) = [P(A) + P(B) + P(C)] - [P(AB) + P(AC) + P(BC)] + P(ABC).}$$

Proof. Replace B by $B \cup C$ in 2-events union formula to get

$$\begin{aligned} P\{A \cup B \cup C\} &= P(A) + P(B \cup C) - P(AB \cup AC) \\ &= P(A) + [P(B) + P(C) - P(BC)] - [P(AB) + P(AC) - P(ABC)] \\ &= \sum P(A) - \sum P(AB) + P(ABC). \end{aligned}$$

Note. For 4 events A_j, $j = 1, 2, 3, 4$ we often write

$S_1 = \sum P(A_j)$, [4 terms], $S_2 = \sum P(A_iA_j)$, [6 terms], $S_3 = \sum P(A_iA_jA_k)$, [4 terms]

$S_4 = P(A_1A_2A_3A_4)$, [one term]. Then

$$P(L_1) = P(\bigcup A_j) = S_1 - S_2 + S_3 - S_4.$$

7. *Probability of Exactly One of the Two Intersecting Events* :

$$\underline{P(A \Delta B) = P(A) + P(B) - 2P(AB).} \qquad [A \Delta B = A\bar{B} \uplus \bar{A}B]$$

Proof.
$$\begin{aligned} P(A \Delta B) &= P(A\bar{B} \uplus \bar{A}B) = P(A\bar{B}) + P(\bar{A}B) && \text{[by } P3' \text{ : finite additivity]} \\ &= [P(A) - P(AB)] + [P(B) - P(AB)] = P(A) + P(B) - 2P(AB). && \text{[by 4]} \end{aligned}$$

Note. Given two events A and B, there occur 12 probabilities :

$P(A), P(\bar{A}), P(B), P(\bar{B}), P(AB), P(\bar{A}B), P(A\bar{B}), P(\bar{A}\bar{B}), P(A \cup B), P(\bar{A} \cup B), P(A \cup \bar{B}), P(\bar{A} \cup \bar{B})$.

However, only 3 can be chosen and the rest can be had through formulae.

1-53. *Poincare's Formula and Inequalities*

With usual notation

$$\boxed{P(L_1) = S_1 - S_2 + S_3 - ... + (-1)^{n+1} S_n} \qquad ...(1)$$

$$(-1)^r P(L_1) \geq (-1)^r [S_1 - S_2 + S_3 - \ldots + (-1)^{r+1} S_r] \qquad \ldots(2)$$

[e.g. $P(L_1) \leq S_1$, $P(L_1) \geq S_1 - S_2$, $P(L_1) \leq S_1 - S_2 + S_3, \ldots$].

Proof. For economy of notation we set

$$P(L_1) = P\{A_1 \cup A_2 \cup \ldots \cup A_n\} = P\{\text{At least one of } A_1, \ldots, A_n \text{ events occurs}\},$$

$$S_r = \sum_{i_1 < i_2 < \ldots < i_r} P(A_{i_1} A_{i_2} \ldots A_{i_r}), 1 \leq r \leq n$$

$$H_r = \sum_{i_1 < \ldots < i_r} I_{i_1} I_{i_2} \ldots I_{i_r} \quad [I_j = \text{Indicator of event } A_j]$$

$$I_{L_1} = 1 - \prod_{j=1}^{n} (1 - I_j) = \begin{cases} 1, & \text{if at least one event } A_j \text{ occur} \\ 0, & \text{otherwise} \end{cases}$$

$$= 1 - \left[1 - \sum_i I_i + \sum_{i<j} I_i I_j - \ldots + (-1)^n I_1 I_2 \ldots I_n\right], \quad \text{[on multiplication]}$$

$$= H_1 - H_2 + H_3 - \ldots + (-1)^{n+1} H_n \qquad \ldots(\text{i})$$

Expected value of this result, using $E(I_{L_1}) = P(L_1)$, $E(H_r) = S_r$ gives

$$\boxed{\mathbf{P(L_1) = S_1 - S_2 + S_3 - \ldots + (-1)^{n+1}. S_n} \text{ (Poincare's formula)}}$$

Now recall the well-known combinatorial identity :

$$\binom{k}{r+1} - \binom{k}{r+2} + \binom{k}{r+3} - \ldots + (-1)^{k-r+1} \binom{k}{k} = \binom{k-1}{r} \geq 0 \qquad \ldots(\text{ii})$$

[Just obtained on equating coefficient of x^{r+1} on both sides of $(1+x)^k (1+x^{-1})^{-1} = x(1+x)^{k-1}$)]

We now transfer r terms from the R.H.S. of (i) to its L.H.S. to get

$$I_{L_1} - H_1 + H_2 - \ldots + (-1)^r H_r = (-1)^r \{H_{r+1} - H_{r+2} + \ldots + (-1)^{n-r+1} H_n\}$$

or $$(-1)^r \{I_{L_1} - H_1 + H_2 - \ldots + (-1)^r H_r\} = H_{r+1} - H_{r+2} + \ldots + (-1)^{n-r+1} H_n \qquad \ldots(\text{iii})$$

Now suppose that exactly k of n events $A_1, \ldots, A_n$ occur. If $k \leq r$, the R.H.S. of (iii) is zero. If $k > r$, the contribution by the terms in R.H.S. of (iii) is

$$\binom{k}{r+1} - \binom{k}{r+2} + \ldots + (-1)^{n-r+1} \binom{k}{k} = \binom{k-1}{r} > 0, \quad \text{[by (ii)]}$$

Hence, no matter how many of the events A_j occur

$$(-1)^r E\{I_{L_1} - H_1 + H_2 - \ldots + (-1)^r H_r\} = E\{H_{r+1} - H_{r+2} + \ldots + (-1)^{n-r+1} H_n\} \geq 0$$

i.e. $$(-1)^r E\{I_{L_1}\} \geq (-1)^r E\{H_1 - H_2 + H_3 - \ldots + (-1)^{r+1} H_r\}$$

$\Rightarrow$ $$(-1)^r P(L_1) \geq (-1)^r [S_1 - S_2 + \ldots + (-1)^{r+1} (S_r].$$

Note. $\underline{S_1 - S_2 \leq P(L_1) \leq S_1.}$ $\underline{\text{[Boole's Inequality]}}$

Extension of Poincare's Formula :

Let L_r denote that *at least* r of the events $A_1, \ldots, A_n$ occur and E_r denote the event that *exactly* r of the events occur. Then

$$P(L_r) = \sum_{k=r}^{n} (-1)^{k-r} \binom{k-1}{r-1} S_k, \quad P(E_r) = \sum_{k=r}^{n} (-1)^{k-r} \binom{k}{r} S_k.$$

Note that $E_r = L_r - L_{r+1}$, so that, $P(E_r) = P(L_r) - P(L_{r+1})$. The following Rider obtains these results when $n = 3$, and $0 \le r \le 3$.

1-54. *Useful Formulas about Three Events*

Let A, B, C be three arbitrary events, and denote

$$S_1 = P(A) + P(B) + P(C), \quad S_2 = P(AB) + P(BC) + P(AC), \quad S_3 = P(ABC).$$

Let E_r, L_r and M_r denote "exactly r events occur", "at least r events occur" and "at most r events occur". Find $P(E_r)$, $P(L_r)$ and $P(M_r)$ for $r = 0, 1, 2, 3$.

Derivations. (a) Firstly, we observe that,

$$P(\bar{A}BC) = P(BC) - P(ABC) = P(BC) - S_3. \qquad \ldots(1)$$

$$P(A\bar{B}\bar{C}) = P(A \cap \overline{B \cup C}) = P(A) - P(AB \cup AC) = P(A) + P(BC) - S_2 + S_3. \qquad \ldots(2)$$

(i) $\quad E_0 = \bar{A}\bar{B}\bar{C}$, so that $P(E_0) = 1 - P(A \cup B \cup C) = 1 - S_1 + S_2 - S_3$.

(ii) $\quad E_1 = A\bar{B}\bar{C} \uplus \bar{A}B\bar{C} \uplus \bar{A}\bar{B}C$

$\therefore \quad P(E_1) = P(A\bar{B}\bar{C}) + P(\bar{A}B\bar{C}) + P(\bar{A}\bar{B}C) = S_1 - 2S_2 + 3S_3$ [by (2)]

(iii) $\quad E_2 = \bar{A}BC \uplus A\bar{B}C \uplus AB\bar{C}$

$\therefore \quad P(E_2) = P(\bar{A}BC) + P(A\bar{B}C) + P(AB\bar{C})$

$= [P(BC) - S_3] + [P(AC) - S_3] + [P(AB) - S_3] = S_2 - 3S_3$. [by (1)]

(iv) $\quad E_3 = ABC$, so that $P(E_3) = P(ABC) = S_3$.

(b) Since $\displaystyle P\{L_k\} = \sum_{j=k}^{3} P(E_j), \qquad P\{M_k\} = \sum_{i=0}^{k} P(E_j)$ $\qquad \ldots(3)$

$$\therefore \quad P(L_0) = \sum_{i=0}^{3} P(E_i) = P(E_0) + \ldots + P(E_3) = 1.$$

$$P(L_1) = \sum_{i=1}^{3} P(E_i) = P(E_1) + P(E_2) + P(E_3) = S_1 - S_2 + S_3$$

$$P(L_2) = \sum_{i=2}^{3} P(E_i) = P(E_2) + P(E_3) = S_2 - 2S_3.$$

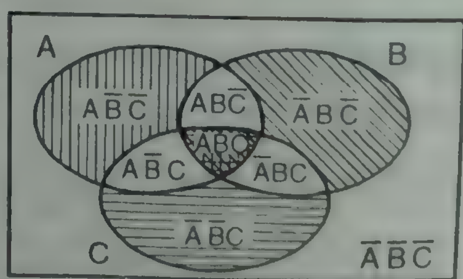
$$P(L_3) = \sum_{i=3}^{3} P(E_i) = P(E_3) = S_3.$$

$$P(M_0) = \sum_{i=0}^{0} P(E_i) = P(E_0) = 1 - S_1 + S_2 - S_3.$$

$$P(M_1) = \sum_{i=0}^{1} P(E_i) = P(E_0) + P(E_1) = 1 - S_2 + 2S_3.$$

$$P(M_2) = \sum_{i=0}^{2} P(E_i) = P(E_0) + P(E_1) + P(E_2) = 1 - S_3.$$

$$P(M_3) = \sum_{i=0}^{3} P(E_i) = P(E_0) + \ldots + P(E_3) = 1.$$



AUBUC

1-55. Some Additional Probability Formulas

1. *Probability of Exactly m of the N Events.* The probability that exactly m unspecified events among the N events $A_1, A_2, \ldots, A_N$ occur simultaneously, written $P_{[m]}$ or $P(E_m)$, is given by

$$P(E_m) = \sum_{k=m}^{N} (-1)^{k-m} \binom{k}{m} S_k, \quad \left[S_r = \sum_{(j_1, \ldots, j_r)} P(A_{j_1} \ldots A_{j_r})\right]. \qquad \ldots(1)$$

Proof. Let $$E_m = \biguplus_{W_m} A_{j_1} A_{j_2} \ldots A_{j_m} \bar{A}_{jm+1} \ldots \bar{A}_{jN}$$

where the disjoint union is over the ${}^N C_m$ possible subsets $W_m = \{j_1, j_2, \ldots, j_m\}$ of size m chosen from the integer set $W = \{1, 2, \ldots, N\}$. The indicator form of the above events, with obvious notation :

$$I(E_m) = \Sigma I_{j_1} I_{j_2} \ldots I_{j_m} (1 - I_{j_{m+1}}) \ldots (1 - I_{j_N}) = \sum_{W_m} I_{j_1} I_{j_2} \ldots I_{j_m} \left\{\sum_{r=0}^{N-m} (-1)^r H_r\right\}.$$

Here $H_r = \Sigma I_{t_1} I_{t_2} \ldots I_{t_r}$, and in it the summation is over all r-elements subsets T_r, of the set $W - W_m$. We rewrite :

$$I(E_m) = \sum_{r=0}^{N-m} (-1)^r \sum_W I_{j_1} I_{j_2} \ldots I_{j_m} \sum_{T_r} I_{t_1} I_{t_2} \ldots I_{t_n} .$$

The set $\{W_m, T_r\} = \{j_1, j_2, \ldots, j_m, t_1, t_2, \ldots, t_r\}$ is just a subset of size $m + r$ of the integer set W since $T_r \cap W_m = \varnothing$. However, some of these sets will be identical because the j's and i's assume all values $1, 2, \ldots, N$. The number of repetitions of a fixed set, like $\{W_m, T_r\}$ is just the number of different ways ${}^{m+1}C_m$ of dividing $\{W_m, T_r\}$ into two sets, containing m and r elements. Hence, reindexing

$$I(E_m) = \sum_{r=0}^{N-m} (-1)^r \binom{m+r}{m} \sum_{W_{r+m}} I_{j_1} I_{j_2} \ldots I_{j_{m+r}} .$$

where $W_{r+m} = \{j_1, j_2, \ldots, j_{r+m}\}$ is a subset of size $r + m$ of the integer set W. Putting $m + r = k$, the above result reads

$$I(E_m) = \sum_{k=m}^{N} (-1)^{k-m} \binom{k}{m} \sum_{W_k} I_{j_1} I_{j_2} \ldots I_{j_k} \qquad \ldots(\text{i})$$

Taking expectations of both sides, using $\mathsf{E}(I_j) = P(A_j)$, etc. we get the stated result :

$$P(E_m) = \sum_{k=m}^{N} (-1)^{k-m} \binom{k}{m} S_k. \qquad \ldots(1)$$

Cor. $$\sum_{r=0}^{N} I_{[r]} x^r = \sum_{s=0}^{N} (x-1)^s J_s,$$

[x : arbitrary variable] where $J_r = \sum_{W_r} I_{j_1} I_{j_2} \ldots I_{j_r} : J_0 = 1.$]

Proof. From Eq. (i), we recover

$$I_{[r]} = \sum_{k=r}^{N} (-1)^{k-r} \binom{k}{r} J_k$$

Multiplying this equation by x^r and summing over r we get

$$\sum_{r=0}^{N} I_{[r]}\, x^r = \sum_{r=0}^{N}\left[\sum_{n=r}^{N}(-1)^{k-r}\binom{k}{r}x^r\right]J_k = \sum_{k=0}^{N}\left[\sum_{r=0}^{k}(-1)^{k-r}\binom{k}{r}x^r\right]J_k \qquad \ldots\text{(ii)}$$

where we have used : $\sum_{i=0}^{N}\sum_{i=j}^{N} a_{ij} = \sum_{i=0}^{N}\sum_{j=0}^{i} a_{ij}$.

Using Binomial expansion we sum up the inner expression in (ii) to get

$$\sum_{r=0}^{N} I_{[r]}\, x^r = \sum_{k=0}^{N}(x-1)^k\binom{k}{r}J_k.$$

2. *Probability of atleast m among the N Events.* The probability $P(m)$ or $P(L_m)$ that at least m $(1 \le m \le N)$ unspecified events of the N events $A_1, A_2, \ldots, A_N$ occur, is given by

$$P(L_m) = \sum_{k=m}^{N}(-1)^{k-m}\binom{k-1}{m-1}S_k. \qquad \ldots(1)$$

Proof. Since $L_m = E_m \uplus E_{m+1} \uplus \ldots \uplus E_N$, hence $P(L_m) = P(E_m) + P(E_{m+1}) + \ldots + P(E_n)$. Using formula for $P(E_r)$, this begets us

$$P(L_m) = \sum_{r=m}^{N}(-1)^{r-m}\binom{r}{m}S_r + \sum_{r=m+1}^{N}(-1)^{r-m-1}\binom{r}{m+1}S_r + \ldots + \sum_{r=N}^{N}(-1)^{r-N}\binom{r}{N}S_r.$$

The coefficient of S_k on the R.H.S., is C_k, where

$$C_k = (-1)^{k-m}\left[\binom{k}{m} - \binom{k}{m+1} + \binom{k}{m+2} - \ldots + (-1)\binom{k}{k}\right].$$

Using Pascal relation : $\binom{k}{a} = \binom{k-1}{a-1} + \binom{k-1}{a}$ [increase head by 1, take greater tail] the above gives

$$C_k = (-1)^{k-m}\left\{\left[\binom{k-1}{m-1} + \binom{k-1}{m}\right] - \left[\binom{k-1}{m} + \binom{k-1}{m+1}\right] + \left[\binom{k-1}{m+1} + \binom{k-1}{m+2}\right] - \ldots\right\}$$

$$= (-1)^{k-m}\binom{k-1}{m-1},$$

all other terms cancel in pairs.

$$\therefore \qquad P(L_m) = \sum_{k=m}^{N} C_k S_k = \sum_{k=m}^{n}(-1)^{k-m}\binom{k-1}{m-1}S_k.$$

3. *Triangle Inequality* : $\underline{P(A \,\Delta\, C) \le P(A \,\Delta\, B) + P(B \,\Delta\, C).}$ $\qquad \ldots(1)$

Proof. From point-set algebra for symmetric differences $A \,\Delta\, C \subseteq (A \,\Delta\, B) \uplus (B \,\Delta\, C)$.

$\therefore \;\; P(A \,\Delta\, C) \le P\{(A \,\Delta\, B) \uplus (B \,\Delta\, C)\} = P(A \,\Delta\, B) + P(B \,\Delta\, C)$

where we have used Monotonic Law (p.18) and finite additivity. This proves the truth of (1).

Note. If we interpret $P(A \,\Delta\, B)$ as the probabilistic distance (metric function) between the events A and B, then relation (1) is the usual triangle inequality.

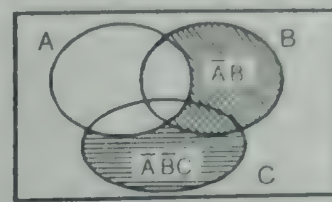
4. Boole's Inequality : If $A_1, A_2, ..., A_n$ are arbitrary events ($A_i \in \mathcal{F}$), then

$$P\{A_1 \cup A_2 \cup ... \cup A_n\} \le P(A_1) + P(A_2) + ... + P(A_n). \qquad ...(1)$$

Case of countably infinite number of events : If $A_1, A_2, ...$ are arbitrary events, then

$$P(A_1 \cup A_2 \cup ... \cup A_n \cup ...) \le P(A_1) + P(A_2) + ... + P(A_n) + ... \qquad ...(2)$$

Proof. We observe that



$$A \cup B \cup C = (A \uplus \bar{A}B \uplus \bar{A}\bar{B}C),$$

and in general

$$A_1 \cup A_2 \cup ... \cup A_n = A_1 \uplus \bar{A}_1 A_2 \uplus \bar{A}_1 \bar{A}_2 A_3 \uplus ... \uplus \bar{A}_1 ... \bar{A}_{n-1} A_n$$

$$\therefore \; P(A_1 \cup A_2 \cup ... \cup A_n) = \sum_{k=1}^{n} P(\bar{A}_1 \bar{A}_2 ... \bar{A}_{k-1} A_k) \le \sum_{k=1}^{n} P(A_k). \qquad \{\bar{A}_1 \bar{A}_2 ... \bar{A}_{k-1} \cdot A_k \subseteq A_k\}$$

If we let $n \to \infty$, the above yields (2).

The result also follows by Mathematical Induction.

Comments. $\text{Max}\{P(A), P(B)\} \le P(A \cup B) \le \min\{1, P(A) + P(B)\}$.

5. Bon-ferroni's Inequality : If $A_1, A_2, ..., A_n$ are arbitrary but compatible events, then

$$P(A_1 A_2, ..., A_n) \ge P(A_1) + P(A_2) + ... + P(A_n) - (n-1). \qquad ...(1)$$

$$P(A_1, A_2, ..., A_n) \ge 1 - P(\bar{A}_1) - P(\bar{A}_2) - ... - P(\bar{A}_n).$$

Proof. We use Principle of Mathematical Induction. For $n = 2$

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 A_2).$$

$$P(A_1 \cup A_2) \le 1 \;\Rightarrow\; P(A_1 A_2) \ge P(A_1) + P(A_1) - (2-1) \qquad ...(i)$$

Thus, (1) is true for $n = 2$, and this starts an induction process. We now assume that (1) is true for $2 \le n = k$; then

$$P(A_1 A_2, ..., A_k) \ge P(A_1) + P(A_2) + ... + P(A_k) - (k-1) \qquad ...(ii)$$

Write $A_1 ... A_k = B_k$, then $A_1 ... A_{k+1} = B_{k+1} = B_k \cap A_{k+1}$, and use (i) to events B_k and A_{k+1} to obtain

$$P(B_k \cap A_{k+1}) \ge P(B_k) + P(A_{k+1}) - (2-1).$$

Using (ii) this becomes : $P(A_1 A_2, ..., A_{k+1}) \ge P(A_1) + ... + P(A_{k+1}) - [(k+1) - 1]$.

This shows that the result is true for $n = k + 1$. As such the process of induction is complete and the result is completely established.

Comments. $\min\{P(A), P(B)\} \ge P(AB) \ge \max\{0, 1 - P(\bar{A}) - P(\bar{B})\}$.

1-56. *Worked-out Problems*

Example 1. A pointer is spun on a fair wheel of chance having its periphery labelled from 0 to 100. Describe the sample space for this experiment, and find the probability that the pointer stops between 45 and 67. What is the probability that the pointer will stop at 93 ?

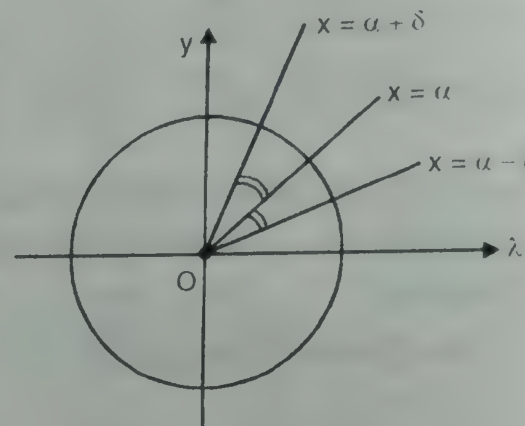
spun - rotate, revolve

Solution. Here sample space is $S = \{0 < X \leq 100\}$. The probability of the pointer falling between any two numbers a, b, $(b \geq a)$ is $p = (b - a)/100$, [because the wheel is fair so probability is uniform], so here $p = (67 - 45)/100 = 0.22$.

Obviously $P(X - 93) = 0$, because the number 93 is only one of infinite number of numbers in sample space S.

Note. The assumption : $p = (b - a)/100$, satisfies the usual axioms. Since $b \geq a$, $p \geq 0$, [$P1$]. When $a = 0$ and $b = 100$, $p = 1$, [$P2$]. By breaking the wheel's periphery into N contiguous segments, $C_n = \{s_{n-1} < s \leq s_n\}$, $\quad s_0 = 0$, $\quad s_n = (n)\ 100/N$, $\quad n = 1, 2, \ldots, N$ we find that $P(C_n) = 1/N$ and thus for any N.

$$P\left(\biguplus_{n=1}^{N} C_n\right) = \sum_{n=1}^{N} P(C_n) = \sum_{n=1}^{N}\left(\frac{1}{N}\right) = 1 = P(S). \qquad [P3]$$



Remarks. For continuous r.v., X, let

$$f(x) = (1/2\pi),\ 0 \leq x \leq 2\pi\ ;\ \text{then}$$

$$P\{\alpha - \delta \leq X \leq \alpha + \delta\} = \int_{\alpha-\delta}^{\alpha+\delta} \frac{dx}{2\pi} = \frac{\delta}{\pi}.$$

$$P\{X = \alpha\} = \lim_{\delta \to 0} P\{X - \alpha \leq X \leq \alpha + \delta\} = 0.$$

However, event $\{X = \alpha\}$ is not impossible; i.e. $P(A) = 0$, even though $A \neq \varnothing$.

Comments. Discrete events can occur on continuous sample spaces even if their probability is zero. The infinite sample space has only one outcome [e] satisfying such a discrete event, so $P\{e\} = 0$, although $\{e\} \neq \varnothing$.

Similarly, events with probability 1 may not occur. For instance, for the wheel of chance the event E = {all numbers occurring except the number x_n} satisfies $P(E) = 1$, although $E \neq S$.

Example 2. Events A and B are such that $A \cup B = S$. Prove

$$P\{\omega \in AB\} = P\{\omega \in A\}\ P\{\omega \in B\} - P\{\omega \notin A\}\ P\{\omega \notin B\}. \qquad \ldots(1)$$

Solution. Write $P(A) = p_1$, $P(B) = p_2$, $P(AB) = p_3$ and then $P(A \cup B) = P(S)$ give

$$P(A) + P(B) - P(AB) = P(S) \quad \Rightarrow \quad p_1 + p_2 = 1 + p_3 \qquad \ldots(2)$$

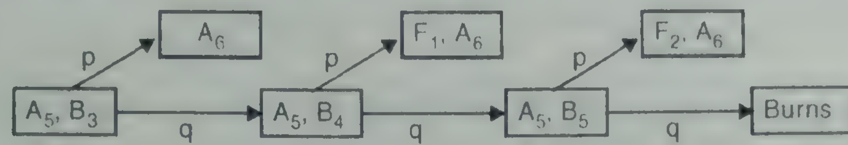
$$P\{\omega \notin A\} = P\{\omega \in \bar{A}B\} = p_2 - p_3$$

$$P\{\omega \notin B\} = P\{\omega \in A\bar{B}\} = p_1 - p_3$$

$$\begin{aligned} \text{R.H.S. of Eqn. (1)} &= p_1 p_2 - (p_2 - p_3)(p_1 - p_3) = p_2 p_3 + p_1 p_3 - p_3^2 \\ &= p_3(p_1 + p_2) - p_3^2 = p_3(1 + p_3) - p_3^2, \qquad \text{[by (2)]} \\ &= p_3 = \text{L.H.S. of Eqn. (1).} \end{aligned}$$

Example 3. A and B play a game in which their chances of winning are p and q, $(p + q = 1)$. The first player to win 6 rounds is the winner and takes the stakes. The game gets interrupted after A has won 5 rounds and B has won 3 rounds. How should the stakes be divided ?

Solution. A tree diagram is suggestive of a good approach.



$$P(W_B) = 1 - P(W_A) = q^3 \qquad \text{(Obvious from tree diagram as well)}$$

The stakes are to be in the ratio : $1 - q^3 : q^3$

[If $p = q = 1/2$, fair game, then stakes are in ratio 7 : 1]

A could win round 6, just after round 5 or after a single failure or after two failures. Hence

$$P(W_A) = P\{A_6 \uplus F_1 A_6 \uplus F_2 A_6\} = P(A_6) + P(F_1)\,P(A_6) + P(F_2)\,P(A_6)$$
$$= p + qp + q^2 p = p(1 + q + q^2) = p(1 - q^3)/(1 - q) = 1 - q^3.$$

***Example* 4.** Three groups of children contain respectively 3 girls, 1 boy ; 2 girls ; 2 boys ; 1 girl and 3 boys. One child is selected at random from each group. Show that the chance that the three selected consist of 1 girl and 2 boys is 13/32.

Solution. Denote the selection of a girl from group k by g_k and that of a boy from group k by b_k. Then

$$E = \{\text{one girl and 2 boys}\} = \{(g_1 b_2 b_3) \uplus (b_1 g_2 b_3) \uplus (b_1\, b_2\, g_3)\}.$$

$$P\{g_1\, b_2\, b_3\} = \frac{3}{4}\cdot\frac{2}{4}\cdot\frac{3}{4} = \frac{9}{32}, \qquad P\{b_1\, g_2\, b_3\} = \frac{1}{4}\cdot\frac{2}{4}\cdot\frac{3}{4} = \frac{3}{32}, \qquad P\{b_1\, b_2\, g_3\} = \frac{1}{4}\cdot\frac{2}{4}\cdot\frac{1}{4} = \frac{1}{32}.$$

$\therefore$ $P\{\text{one girl and 2 boys}\} = P\{g_1\, b_2\, b_3\} + P\{b_1\, g_2\, b_3\} + P\{b_1\, b_2\, g_2\} = 13/32.$

***Example* 5.** Show that Boole's inequality and the Bonferroni's inequality are essentially the same thing.

Solution. Boole's inequality applied to $\overline{A}_1, \overline{A}_2, ..., \overline{A}_n$ yields

$$P(\overline{A}_1 \cup \overline{A}_2 \cup ... \cup \overline{A}_n) \le P(\overline{A}_1) + P(\overline{A}_2) + ... + P(\overline{A}_n) \qquad ...(1)$$

Since $\overline{A}_1 \cup \overline{A}_2 \cup ... \cup \overline{A}_n = (A_1\, A_2\, ...\, A_n)^c$, and $P(\overline{B}) = 1 - P(B)$, (1) yields by Negation rule

$1 - P\,(A_1, A_2, ..., A_n) \le n - \{P(A_1) + P(A_2) + ... + P(A_n)\}$

i.e. $P(A_1, A_2, ..., A_n\} \ge \{P(A_1) + P(A_2) + ... + P(A_n)\} - (n - 1).$ ...(2)

This is Bonferron's inequality : it allows to bound the probability of simultaneous event (i.e. intersection) in terms of probability of the individual events.

***Example* 6.** Ten tickets are numbered 1, ..., N respectively, n tickets are selected, one at a time, with replacement. Find the probability that the highest number appearing on a selected ticket is m.

Solution. Observe : $P(X = m) = P\{(m - 1) < X \le m\}$

$$= P(X \le m) - P(X \le m - 1) \quad \text{[Difference method]} \quad ...(1)$$

Now, the probability that one ticket has a number $\le m$ is $m/N = p$ (say). It follows by independence [selections with replacement that the probability of all 11 tickets chosen have number $\le m$ is p^n. This means the probability that the highest number on n selected

tickets is $\leq m$ is $p^5 = (m/10)^5$. By the same logic, the probability that the highest number on n selected tickets is $\leq (m-1)$ is $[(m-1)/10]^5$. Using (1)

$$P(X = m) = \left(\frac{m}{N}\right)^n - \left(\frac{m-1}{N}\right)^n = \frac{(m)^n - (m-1)^n}{(N)^n}.$$

Problems with Solutions Provided at the End of the Text

1*. Prove or disprove the following :

(a) If $P(A) = 1/3$, $P(\bar{B}) = 1/4$, then $A \cap B = \varnothing$. (b) If $P(A) = 0$, then $P(AB) = 0$.

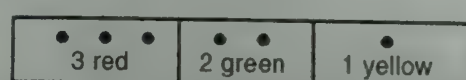
(c) If $P(A) = P(B) = p$, then $P(AB) \leq p^2$. (d) If $P(A) = P(\bar{B})$, then $\bar{A} = B$.

2*. A box contains 10 balls numbered 1 through 10. Five balls are drawn successively at random and without replacement. Let A be the event that exactly two odd-numbered balls are drawn and that they occur on odd-numbered draws from the box. Evaluate $P(A)$.

3*. Four cards are drawn out of a well-shuffled deck of 52 cards. What is the probability of getting exactly two spades and exactly two aces ?

4*. Three urns contain respectively 1 white, 2 black balls; 3 white, 1 black balls; 2 white, 3 black balls. One ball is taken from each urn. What is the probability that among the balls drawn, there are 2 white and 1 black balls ?

5*. An urn contains 3 red balls, 2 green balls and 1 yellow ball. These balls are selected at random and without replacement from the urn. What is the probability that at least one colour is **not** drawn ?



6*. A box contains 3 white and 3 black balls and a person draws out three balls at random. He then drops three yellow balls into the box and after shuffling draws out again three balls. Find the chance that the latter three balls are of different colours.

7*. A bag contains 4 white, 5 red and 6 black balls. Four balls are drawn at random. Find the probability that

(a) No ball drawn is black. (b) Exactly two are black.

(c) All are of the same colour. (d) At least one ball of each colour.

8*. What is the probability of getting a *void* in at least one suit in bridge ? (That is to say, not possessing any cards of at least one suit).

9*. n cards are drawn from an ordinary pack, with replacement. What is the probability that each of the four suits will be represented at least once among the n cards.

10*. n balls are randomly distributed among three boxes. Find the probabilities of :

(a) No empty boxes, (b) Exactly one empty box, (c) Exactly two empty boxes.

11*. Arrange the following in increasing order of magnitude :

$$P(A),\ P(A \cup B),\ P(A \cap B),\ P(A) + P(B).$$

12*. (Implication Rule). Let A, B, C be three events and assume that the simultaneous occurrence of A and B implies occurrence of C; then

$$P(C) \geq P(A) + P(B) - 1, \text{ or } P(\bar{C}) \leq P(\bar{A}) + P(\bar{B}).$$

13*. If $P(A) = 0.9$, $P(B) = 0.8$. Show that $P(AB) \geq 0.7$.

14*. n pairs of shoes are in a closet but the $2n$ individual shoes are thoroughly mixed. r shoes are selected at random. Find the probability that there is at least one pair among the r selected.

15*. In a lottery m tickets are drawn at a time out of the total number of n tickets, and returned before the next drawing is made. Show that the chance that in k drawings, each of the numbers, 1, 2, 3, ..., n, will appear at least once, is given by

$$P_k = 1 - \binom{n}{1}\left(1-\frac{m}{n}\right)^k + \binom{n}{2}\left(1-\frac{m}{n}\right)^k\left(1-\frac{m}{n-1}\right)^k + \binom{n}{3}\left(1-\frac{m}{n}\right)^k\left(1-\frac{m}{n-1}\right)^k\left(1-\frac{m}{n-2}\right)^k + \ldots$$

Exercise 1(c)

1. If $d(A, B) = P(A \Delta B)$, show that d has all the properties of a metric function.
2. Write $A \cup B \cup C$ as a union of disjoint sets and evaluate $P(A \cup B \cup C)$. Also establish :
$$P(ABC) \leq P(AB) \leq (P(A \cup B) \leq P(A \cup B \cup C) \leq P(A) + P(B) + P(C).$$
3. (a) If A flips n fair coins and B flips $n + 1$ fair coins, show that the probability that B flips more heads than A is 1/2.

   (b) Show that the chance of drawing 7 or 11 with two dice, each having six faces numbered 1 to 6 is 2/9.
4. Let $\Omega = \{1, 1/2, (1/2)^2, ..., (1/2)^n\}$, n fixed and $A = \{1, 1/2\}$, $B = \{\omega : \omega\ (1/2)^j, j \geq 1 \text{ and even}\}$, with $P(A) = 1/8$, $P(B) = 1/2$. Show that $P(A \cup B) = 5/8$ and $P(\bar{A}\,\bar{B}) = 3/8$.
5. Let $S = \{e_1, e_2, ..., e_6\}$ with $p_i = P\{e_i\}$ assigned as :

   $p_1 = 0.05$, $p_2 = 0.25$, $p_3 = 0.10$,
   $p_4 = 0.20$, $p_5 = 0.30$, $p_6 = 0.10$.

   Compute the probabilities of the events
   $A = \{e_1, e_2, e_3\}$, $B = \{e_2, e_3, e_4, e_5\}$, $C = \{e_1, e_6\}$, $D = \{e_6\}$, $A \cap B$, $(A \cap C) \cup B$, $(A \cup B) \cap C$, $(A \cap D) \cup (B \cup C)$ [**Ans.** 0.40, 0.85, 0.15, 0.10, 0.35, 0.90, 0.05, 1]
6. A, B, C are disjoint events associated with an experiment $\{A, B, C\}$. If $P(B) = \frac{3}{2} P(A)$ and $P(C) = \frac{1}{2} P(B)$, find $P(A)$, $P(C)$, and $P(A \cup C)$. [**Ans.** $4p$, $3p$, $7p$, $p = 1/3$]
7. Let $\Omega = \{x : 0 < x < 10\}$ be a sample space and define events $A_1 = \{x : 0 < x \leq 2\}$, $A_2 = \{x : 0 < x \leq 4\}$, $A_3 = \{x : 0 < x < 2) \cup \{x : 4 < x < 10\}$. If $P(A_1) = 3/8$, $P(A_2) = 7/8$, show that $P(A_3) = 0.5$.
8. (a) Events A and B are such that $P(A) = 3/4$ and $P(B) = 5/8$, Show that

   (i) $P(A \cup B) \geq 3/4$, (ii) $3/8 \leq P(AB) \leq 5/8$, (iii) $1/8 \leq P(A\bar{B}) \leq 3/8$.

   (b) Given $P(A) = 3/4$, $P(B) = 3/8$. Show that $P(A \cup B) \geq 3/4$; $1/8 \leq P(AB) \leq 3/8$.
   Given $P(A) = 1/3$, $P(B) = 1/4$. Show that $P(A \cup B) \geq 1/3$; $0 \leq P(AB) \leq 1/4$.
9. Events A and B are such that $P(A \cup B) = 3/4$, $P(AB) = 1/4$ and $P(\bar{A}) = 2/3$. Show that $P(B) = 2/3$, and $P(A\bar{B}) = 1/12$.

10. Examine the consistency of the following data :
$P(A) = 0.4$, $P(B) = 0.3$, $P(C) = 0.7$; $P(AB) = 0.2$, $P(BC) = 0.2$, $P(AC) = 0.4$, $P(\overline{A}\overline{B}\overline{C}) = 0.5$.
[**Ans.** Inconsistent]

11. Defects are classified as of a type A, B, C and the following probabilities have been determined from available production data :
$P(A) = 0.20$, $P(B) = 0.16$, $P(C) = 0.14$, $P(AB) = 0.08$, $P(AC) = 0.05$, $P(BC) = 0.04$, $P(ABC) = 0.02$.
What is the probability that a randomly selected item of product will exhibit at least one type of defect ? Show that probability that it exhibits both A and B type defects, but is free from type C defect is 0.06.

12. Which of the following are probability spaces and why ?
(i) $\Omega\{a, b\}$. $\mathcal{F} = \{(a), (b)\}$, $P\{a\} = P\{b\} = 1/2$.
(ii) $\Omega\{1, 2\}$, $\mathcal{F} = \{\varnothing, \{1\}, \{2\}, \Omega\}$, $P(\varnothing) = P\{1\} = P\{2\} = 0$, $P(\Omega) = 1$.
(iii) $\Omega = \{a\}$, $\mathcal{F} = \{\varnothing, \Omega\}$, $P\{\varnothing\} = 0$, $P(\Omega\} = 1$.

13. (a) A sample space S consists of three elementary events e_1, e_2, e_3. Find for each of the following whether it is a probability model :
(i) $P(e_1) = 0.3$, $P(e_2) = 0.6$, $P(e_3) = 0.1$ [Yes]
(ii) $P(e_1) = 0.6$, $P(e_2) = 0.9$, $P(e_3) = -0.5$. [No]

(b) For disjoint events A, B, C show that the following are not permissible assignments of probabilities
$$P(A) = 2/3,\ P(B) = 1/4,\ P(C) = 1/6.$$

(c) Which of the following are probability measures on the events of $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$?

Event	$\varnothing$	$\{\omega_1\}$	$\{\omega_2\}$	$\{\omega_3\}$	$\{\omega_1, \omega_2\}$	$\{\omega_1, \omega_3\}$	$\{\omega_2, \omega_3\}$	Ω
P_1	0	$\frac{1}{2}$	0	$\frac{1}{2}$	$\frac{1}{2}$	1	$\frac{1}{2}$	1
P_2	0	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{5}{4}$	$\frac{3}{4}$	$\frac{1}{2}$	1

14. (a) A card is drawn from a well-shuffled pack of playing cards. Show that the probability that it is either a spade or an ace is 4/13.

(b) Two cards are randomly extracted from a deck. Prove that $P\{$2 cards form a blackjack$\}$ = 32/663, [blackjack = an ace and an other card as ten,jack, queen or king].

15. A card is drawn at random from a full pack. If A and B denote getting a heart card and getting a face card, show that $P(A)$, $P(A\overline{B})$, $P(A \cup B)$, $P(AB)$, $P(\overline{A} \cup B)$, $P(\overline{A}B)$, $P(A \cup \overline{B})$ are respectively equal to $13p$, $10p$, $22p$, $3p$, $42p$, $9p$, $43p$ where $p = 1/52$.

16. (a) Two dice are rolled. If A denotes the event of getting odd sum of the points on the upturned faces and B denotes the event that there is at least one 3 shown, show that $P(A \cup B)$, $P(AB)$, $P(A - B)$ and $P(A\overline{B} \cup \overline{A})$ are respectively equal to 23/36, 1/6, 1/3, 5/6.

(b) Coin C_1 comes up heads with probability p_1 and coin C_2 comes up heads with probability $p_2 > p_1$. You win a bet if in 3 tosses you get at least 2 heads in succession. You are to toss the coins alternately starting with either coin. What coin would you select to start the game ?

17. From a group of 8 children, 5 boys and 3 girls, three children are chosen at random. Show that the probabilities that the elected group contains :
(a) no girl, (b) only one girl, (c) one particular girl, (d) at least one girl, (*e*) more girls than boys, are respectively $10p$. $30p$, $21p$, $46p$, $16p$ where $p = 1/56$.

18. A person has to pass through 4 traffic lights. Let X denote the number of red lights. Let $P(X=0)=0.06$, $P(X=1)=0.24$, $P(X=2)=0.35$, $P(X=3)=0.25$, $P(X=4)=0.10$. Show that $P(X \geq 2)=0.70$, $P(X>2)=0.35$, $P(X \leq 2)=0.65$.

19. An urn A contains four tickets numbered 1, 2, 3, 4 and the urn B contains six tickets numbered 2, 4, 6, 7, 8, 9. One of the two urns is chosen at random and a ticket is drawn at random from it. Show that the probability that the ticket drawn bears the number (a) 2 or 4 is 5/12, (b) 3 is 1/8 and (c) 1 or 9 is 5/24.

20. If A, B, C are three events and if $P(A)=0.40$, $P(B)=0.55$, $P(C)=0.50$, $P(BC)=0.25$, $P(CA)=0.20$, $P(AB)=0.15$, $P(ABC)=0.05$, find $P(E_r)$, $P(L_r)$, $P(M_r)$, $r=1, 2, 3$; where E_r, L_r, M_r denote exactly r, at least r and atmost r events to occur.

[**Ans.** $P(E_r)$ = 0.40, 0.45, 0.50, $P(L_r)$ = 0.90, 0.50, 0.05, $P(M_r)$ = 0.50, 0.95, 1]

21. (a) Five cards are drawn from a pack of 52. Show that the chance that these 5 will contain (i) just one king is 0.2995, (ii) at least one king is 0.3412.

(b) Cards are drawn successively from an ordinary deck, until an ace appears. Find the chance that an ace appears (i) at the kth draw, (ii) after the kth card.

22. Show that $|P(A)-P(B)| \leq P(A\bar{B} \cup \bar{A}B)$.

If $P(A)=0.4$, $P(B)=0.7$, determine the maximum and minimum possible values of $P(AB)$ and the conditions under which each of these values is attained. [**Ans.** $0.1 \leq P(AB) \leq 0.4$]

23. Suppose that each of n sticks is broken into one long and one short part. The $2n$ parts are then shuffled and arranged into n pairs from which new sticks are formed. Show that

(a) P\{All parts are joined in original order\} $= n!\, 2^n/(2n)!$.

(b) P\{All long parts are paired with short parts\} $= (n!)^2\, 2^n/(2n)!$.

(c) P\{At least one of the original sticks is formed $= \Sigma(-1)^{k-1}\binom{n}{k} 2^k . (2n-2k)!/(2n)!$.

24. An event A of the sample space Ω is called P-null if $P(A)=0$, and P-sure if $P(A)=1$. Establish.

(a) If both A and B are P-null, so are AB, $A \cup B$ and $A\bar{B}$.

If A is P-null and B is any event, then, $P(AB)=0$ and $P(A \cup B)=P(B)$.

(b) If both A and B are P-sure, so are AB, $A \cup B$ and $A \cup \bar{B}$.

If A is P-sure and B is any event, then $P(A \cup B)=1$ and $P(AB)=P(B)$.

25. Find $P(A \cap C)$ if $P(\bar{A}\,\bar{B}\,\bar{C})=0.1$, $P(A \Delta B \Delta C)=0.6$, $P[B \cap (A \cup C)]=0.3$, $P(ABC)=0.1$. [**Ans.** 0.2]

26. A die is thrown N times. Show that, for $1 \leq k \leq 6$,

$$P[\text{Each of } k \text{ given faces appears atleast once}\} = \sum_{r=0}^{k} (-1)^k \binom{k}{r}\left(1-\frac{r}{6}\right)^N.$$

27. Let $(\Omega, \mathcal{F})$ be a measurable space and let $P_1, P_2, \ldots, P_n$ be a sequence (collection) of probability measures defined on $\mathcal{F}$. If $a_i \geq 0$, $i=1, 2, \ldots, n$ are real numbers such that $\Sigma a_i = 1$, then the function P defined on $\mathcal{F}$ by the relation.

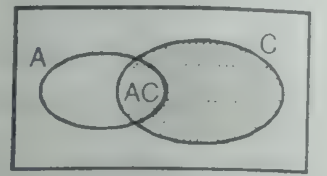
(a) $P(A)=\Sigma a_i P_i(A)$, (b) $P(A)=\Sigma(1/2)^n P_n(A)$, $n=1, 2, \ldots, \infty$

is a probability measures on $\mathcal{F}$.

28. Suppose that the probabilities $P(A_1)$, $P(A_2)$, ..., $P(A_n)$ are known and that $A_1, A_2, \ldots, A_n$ form a partition of Ω. Show that it is possible to calculate the probability of any member of $\mathcal{F}$ (the field of events).

1-60. *Conditional Probability*

Let C be an arbitrary event in a sample space Ω with $P(C) > 0$. The probability that an event A occurs once C has occurred, or, in other words, the conditional probability of A given C, notated $P(A \mid C)$, is defined by :



Reduced sample space

$$P(A \mid C) = P(A \cap C)/P(C).$$

In a certain sense, $P(A \mid C)$ measures the relative probability of A w.r.t. the new sample space (the reduced space) C. (See Fig.). If Ω is an equiprobable space, then

$$P(A \cap C) = n(AC)/n(\Omega), \quad P(C) = n(C)/n(\Omega).$$

$$\therefore \quad P(A \mid C) = n(AC)/n(C) = \frac{\text{Number of ways both } A \text{ and } C \text{ can occur}}{\text{Number of ways } C \text{ can occur}}.$$

Example. A family has two children. Find the probability p that both children are boys if it is known that : (i) One of the children is a boy, (ii) The older child is a boy.

Solution. Here $\Omega = \{bb, bg, gb, gg\}$. The reduced space is

(i) $C = \{bb, bg, gb\}$, so $p = 1/3$. (ii) $C = \{bb, bg\}$, so $p = 2/4 = 1/2$.

1-61. *Conditional Probability as a Genuine Probability Measure*

To show that the conditional probability function $P(* \mid C)$ satisfies the three axioms of Probability Theory.

$[P_1]$ $0 \leq P(A \mid C) \leq 1$; $[P_2]$ $P(S \mid C) = 1$

$[P_3]$ $P\{(A_1 \uplus A_2 \uplus \ldots) \mid C\} = P(A_1 \mid C) + P\{A_2(C)\} + \ldots$

Proof. (1) Since $A \cap C \subset C$, we get $P(AC) \leq P(C)$. Thus,

$P(A \mid C) = P(AC)/P(C) \leq 1$, and is also non-negative because $P(AC) > 0$ and $P(C) > 0$.

(2) Since $S \cap C = C$, $P(S \mid C) = P(S \cap C)/P(C) = P(C)/P(C) = 1$.

(3) If $A_i \cap A_j = \varnothing$ $(i \neq j)$, then $(A_i \cap C) \cap (A_j \cap C) = \varnothing$, so A_iC and A_jC are disjoint events. Now,

$$\therefore \quad P\{[A_1 \uplus A_2 \uplus \ldots] \cap C\} = P\{A_1C \uplus A_2C \uplus \ldots\} = P(A_1C) + P(A_2C) + \ldots$$

i.e. $$\frac{P[(A_1 \uplus A_2 \uplus \ldots) \cap C]}{P(C)} = \frac{P(A_1C)}{P(C)} + \frac{P(A_2C)}{P(C)} + \ldots \qquad \left[\because \frac{P(D \cap C)}{P(C)} = P(D \mid C)\right]$$

or $$P[A_1 \uplus A_2 \uplus \ldots) \mid C] = P(A_1 \mid C) + P(A_2 \mid C) + \ldots$$

Comments. Since $P(* \mid C)$ is a non-negative normed measure and is also σ-additive, it follows that $P(* \mid C)$ must satisfy any consequence of the usual probability axioms ; e.g.

$$P(A \cup B \mid C) = P(A \mid C) + P(B \mid C) - P(AB \mid C), \; P(A \mid C) \leq P(B \mid C); \text{ if } A \subseteq B, \text{ etc.}$$

Reversal Identity : $$\boxed{P(B \mid A) = \frac{P(A \mid B) . P(B)}{P(A)}}$$

As $$P(B \mid A) = \frac{P(AB)}{P(A)}, \quad P(A \mid B) = \frac{P(AB)}{P(B)}$$

$$\therefore \quad P(AB) = P(B \mid A) . P(A) = P(A \mid B) . P(B).$$

Thus Product Rule yields the Reversal Identity readily.

1-62. General Product Rule

For arbitrary events $A_1, A_2, \ldots, A_n$

$$P(A_1, A_2, \ldots, A_n) = P(A_1)\, P(A_2 \mid A_1)\, P(A_3 \mid A_1 A_2) \ldots P(A_n \mid A_1, A_2, \ldots, A_{n-1}). \quad \ldots(1)$$

Proof. Since $A_1 \supseteq A_1 A_2 \supseteq \ldots \supseteq A_1, A_2, \ldots, A_n$, we readily obtain, by Monotonicity :

$$P(A_1) \geq P(A_1 A_2) \geq \ldots \geq P(A_1, A_2, \ldots, A_n) > 0.$$

It follows that all the conditional probabilities in (1) are well defined. We now use the method of mathemetical induction. The result is true for $n = 2$ [trivially from the definition of conditional probability]. We assume that the result is true for $2 \leq n = m$; hence

$$P(A_1, A_2, \ldots, A_m) = P(A_1)\, P(A_2 \mid A_1)\, P(A_3 \mid A_1 A_2) \ldots P(A_m \mid A_1, A_2, \ldots, A_{m-1}) \quad \ldots(\text{i})$$

From : $\quad P(BA_{m+1}) = P(B)\, P(A_{m+1} \mid B)$, with $B = A_1 A_2 \ldots A_m$, we get

$$P[A_1, A_2, \ldots, A_m) A_{m+1}] = P(A_1, A_2, \ldots, A_m)\, P(A_{m+1} \mid A_1, A_2, \ldots, A_m) \quad \ldots(\text{ii})$$

$$= P(A_1)\, P(A_2 \mid A_1) \ldots P(A_m \mid A_1 A_2 \ldots A_{m-1})$$

$$\times P(A_{m+1} \mid A_1, A_2, \ldots, A_m)$$

where (i) is inserted into (ii). This shows that result (1) is true for $n = m + 1$, if it is true for $n = m$. By Induction Principle, (1) is true for all integral values of n.

Aliter. $P(A_1 A_2 \ldots A_n) = P(A_1) . \dfrac{P(A_1 A_2)}{P(A_1)} . \dfrac{P(A_1 A_2 A_3)}{P(A_1 A_2)} \cdots \dfrac{P(A_1 A_2 \ldots A_{n-1})}{P(A_1 A_2 \ldots A_{n-2})} . \dfrac{P(A_1 A_2 \ldots A_{n-1} A_n)}{P(A_1 A_2 \ldots A_{n-1})}$

$$= P(A_1) . P(A_2 \mid A_1) . P(A_3 \mid A_1 A_2) \ldots P(A_n \mid A_1 A_2 \ldots A_{n-1}). \quad \text{[Def. Cond. Probability]}$$

Cor. If events are independent, $P(A_i \mid A_j) = P(A_i)$, $P(A_i \mid A_j A_k) = P(A_i)$; etc. and (1) translates to

$$P(A_1 A_2 \ldots A_n) = P(A_1)\, P(A_2) \ldots P(A_n). \quad \ldots(2)$$

Product Rule for Conditional Probability :

$$P(AB \mid C) = P(A \mid C)\, P(B \mid AC). \quad \text{[Proof trivial]}$$

$$P(A_1 A_2 \ldots A_n \mid C) = P(A_1 \mid C)\, P(A_2 \mid A_1 C)\, P(A_3 \mid A_1 A_2 C) \ldots P(A_n \mid A_1 A_2 \ldots A_{n-1} C).$$

1-63. Worked-out Problems

***Example* 1.** In terms of probabilities : $p_1 = P(A)$, $p_2 = P(B)$, $p_3 = P(AB)$, $[p_1\, p_2\, p_3 > 0]$, express the following in terms of p_1, p_2, p_3 :

(a) $P(\overline{A \cup B})$	(b) $P(\bar{A} \cup \bar{B})$	(c) $P(\bar{A} B)$
(d) $P(\bar{A} \cup B)$	(e) $P(\bar{A} \bar{B})$	(f) $P(A \bar{B})$
(g) $P(A \mid B)$	(h) $P(B \mid \bar{A})$	(i) $P[\bar{A} \cap (A \cup B)]$
(j) $P[A \cup (\bar{A} \cap B)]$	(k) $P(\bar{A} \mid \bar{B})$.	

Solution. It is useful to refer to Euler's diagram [p. 3]

(a) $P(\overline{A \cup B}) = 1 - P(A \cup B) = 1 - [P(A) + P(B) - P(AB)] = 1 - p_1 - p_2 + p_3.$

(b) $P(\overline{A} \cup \overline{B}) = P(\overline{A \cap B}) = 1 - P(A \cap B) = 1 - p_3.$

(c) $P(\overline{A}\, B) = P(B - AB) = P(B) - P(AB) = p_2 - p_3.$

(d) $P(\overline{A} \cup B) = P(\overline{A}) + P(B) - P(\overline{A}B) = 1 - p_1 + p_2 - (p_2 - p_3) = 1 - p_1 + p_3.$ [by (c)]

(e) $P(\overline{A} \cap \overline{B}) = P(\overline{A \cup B}) = 1 - p_1 - p_2 + p_3.$ [by (a)]

(f) $P(A\overline{B}) = P(A) - P(AB) = p_1 - p_3.$

(g) $P(A \mid B) = P(AB)/P(B) = p_3/p_2.$

(h) $P(B \mid \overline{A}) = P(\overline{A}\, B)/P(\overline{A}) = (p_2 - p_3)/(1 - p_1).$ [by (c)]

(i) $P[\overline{A} \cap (A \cup B)] = P[\overline{A} \cap A) \cup (\overline{A}\, B)] = P(\varnothing \cup \overline{A}\, B) = P(\overline{A}\, B) = p_2 - p_3.$ [by (c)]

(j) $P[A \cup (\overline{A} \cap B)] = P[(A \cup B) = P(A) + P(B) - P(AB) = p_1 + p_2 - p_3.$

(k) $P(\overline{A} \mid \overline{B}) = P(\overline{A}\, \overline{B})/P(\overline{B}) = (1 - p_1 - p_2 + p_3)/(1 - p_2).$ [by (*e*)]

Definitions. If A and B are disjoint, $p_3 = P(AB) = 0$.

If A and B are independent, $P(AB) = P(A)\, P(B) \Rightarrow p_3 = p_1\, p_2.$

Several of the above results can then be simplified by these additional conditions.

Example 2. It is felt that probabilities are 0.20, 0.40, 0.30 and 0.10 that teams of four zones A, B, C, D will win their final championship. For some reason, zone B is declared ineligible for championship. Find the probability that zone A will win the championship.

Solution. The given data reads

$$P(A) = 0.2,\ P(B) = 0.4,\ P(C) = 0.3,\ P(D) = 0.1,\ P(\overline{B}) = 0.6.$$

We need find $P(A \mid \overline{B})$. Now

$$P(A \mid \overline{B}) = \frac{P(A\overline{B})}{P(\overline{B})} = \frac{P(A) - P(AB)}{P(\overline{B})} = \frac{P(A)}{P(\overline{B})} = \frac{0.2}{0.6} = \frac{1}{3}. \qquad [P(AB) = P(\varnothing) = 0]$$

Note. Since $A \subset \overline{B}, C \subset \overline{B}, D \subset \overline{B}$, we find $A\overline{B} = A, C\overline{B} = C$ and $D\overline{B} = D.$ Hence

$$P(A \mid \overline{B}) = P(A\overline{B})/P(\overline{B}) = P(A)/P(\overline{B}) = \tfrac{1}{3}.$$

Example 3. If $P(A) = a$, $P(B) = b$, then $P(A \mid B) \geq (a + b - 1)/b$.

Solution. From $P(A \cup B) \leq 1$ we obtain : $P(A) + P(B) - P(AB) \leq 1.$ ...(i)

Since $P(AB) = P(B)\, P(A \mid B) = bP\,(A \mid B)$, relation (i) reduces to

$a + b - b\, P(A \mid B) \leq 1 \Rightarrow P(A \mid B) \geq (a + b - 1)/b.$

Example 4. If A and B are two events, prove that

$$P(A \mid \overline{B}) = [P(A) - P(AB)]\,/\,[1 - P(B)]. \qquad [P(B) \neq 1].$$

Deduce that $P(AB) \geq P(A) + P(B) - 1$. Also show that $P(A) > P(A \mid B)$ if $P(A \mid \overline{B}) > P(A)$.

Solution. (i) By definition,

$$P(A \mid \overline{B}) = P(A\overline{B})/P(\overline{B}) = [P(A) - P(AB)]/[1 - P(B)]. \qquad ...(1)$$

Using $P(A|\overline{B}) \le 1$ provides

$P(A) - P(AB) \le 1 - P(B)$ or $P(AB) \ge P(A) + P(B) - 1$.

(ii) $P(A|\overline{B}) > P(A) \Leftrightarrow P(A) - P(AB) > P(A)[1 - P(B)]$ [by (1)]

$$\Leftrightarrow P(AB) < P(A)\,P(B) \Leftrightarrow P(A|B) < P(A).$$

Similarly, $P(A|\overline{B}) < P(A) \Leftrightarrow P(A|B) > P(A)$.

***Example* 5.** Two cards are drawn at random, without replacement, from an ordinary deck. Find the probability of drawing either an ace or a spade.

Solution. Write A_S = Ace of spade, A_S' = Ace of non-spades and define events.

A = {The duo contains ace or aces} = $\{A_1\overline{A}_2 \uplus A_1A_2\}$, A_j for aces, etc.

S = {The duo contains spade or spades} = $\{S_1\overline{S}_2 \uplus S_1S_2\}$, S_j for spades, etc.

$A \cap S$ = {The duo contains an ace and a spade} = $\{(A_SC_2 \uplus C_1A_S) \uplus (S_1\, A_S' \uplus A_S'\, S_2)\}$ where C_1, C_2 are any cards, S_1, S_2, are spade cards.

$$P(A) = P(A_1\,\overline{A}_2) + P(A_1\,A_2) = \left\{\binom{4}{1}\binom{48}{1}\Big/\binom{52}{2}\right\} + \left\{\binom{4}{2}\Big/\binom{52}{2}\right\} = \frac{33}{221} = 0.1493.$$

$$P(S) = P(S_1\,\overline{S}_2) + P(S_1\,S_2) = \left\{\binom{13}{1}\binom{39}{1}\Big/\binom{52}{2}\right\} + \left\{\binom{13}{2}\Big/\binom{52}{2}\right\} = \frac{15}{34} = 0.4412.$$

$P(AS) = P(A_SC_2 \uplus C_1A_S) + P(S_1A_S' \uplus A_S'\,S_2\} = 2P(A_SC_2) + 2P(A_S'\,S_2)$, by symmetry

$$= 2.\frac{1}{52}.\frac{51}{51} + 2.\frac{3}{52}.\frac{12}{51} = \left(\frac{29}{442}\right) = 0.0656. \qquad (S_2 \text{ has to contain no ace})$$

$\therefore$ $P(A \cup S) = P(A) + P(S) - P(AS) = 0.1493 + 0.4412 - 0.0656 = 0.5249$.

***Example* 6.** What is the probability that (i) no two, (ii) at least two, out of a gathering of n people ($n \le 365$) have the same birth date ?

Solution. We assume a non-leap year so that $\Omega = \{1, 2, 3, ..., 365\}$. Define the events

A_j = {jth persons's birthday differs from all other person's birthday}, $j = 1, 2, ...,$

N_2 = {No two people have the same birthday} = $A_1A_2 \ldots A_n$. (All with different birthdays)

Assuming equi-likelihood, writing $\lambda = 1/365$, we have

$P(A_1) = 365/365 = 1$, $P(A_2 | A_1) = 364/365) = (1 - \lambda)$, $P(A_2 | A_1A_2) = 363/365 = (1 - 2\lambda)$.

$P(A_j | A_1A_2 \ldots A_{j-1}) = (365 - j + 1)/365 = [1 - (j - 1)\,\lambda]$.

$$\therefore \quad P(N_2) = P(A_1, A_2, \ldots, A_n) = P(A_1)\,P(A_2 | A_1) \ldots P(A_n | A_1, A_2, \ldots, A_{n-1})$$

$$= 1\,.\,(1 - \lambda)\,(1 - 2\lambda) \ldots [1 - (n - 1)\,\lambda] = 365\,.\,364 \ldots (366 - n)/(365)^n.$$

$$P(L_2) = P\{\text{At least 2 people have the same birthday}\} = 1 - P(N_2).$$

Remarks. If $n = 26$, then using approximations, $\log(1 + x) = x$, we get

$$P(N_2) = (1 - \lambda)\,(1 - 2\lambda)\,(1 - 3\lambda) \ldots (1 - 25\lambda)$$

$$\therefore \qquad \log P(N_2) = \sum_{j=1}^{25} \ln(1-\lambda_j) = \sum_{j=1}^{25} \frac{(-j)}{365} = -\frac{1}{365}\left(\frac{1+25}{2}\right).25 = -\frac{65}{72} = -0.89.$$

$$\therefore \qquad P(N_2) = e^{-0.89} = 0.412.$$

Thus, in a gathering of 26 poeple, the probability is 0.412 that no two of them have the same birthday.

When $\quad n = 50,\ P(N_2) \approx 0.03,\ P(L_2) \approx 0.97$

When $\quad n = 23,\ P(N_2) \approx 0.493,\ P(L_2) \approx 0.507$

Thus, in a gathering of 23 persons, the probability that at least two of them have the same birthday exceeds 0.5.

***Example* 7.** Cards are dealt one by one from a well-shuffled deck until an ace appears. Show that the chance that exactly n cards are dealt in all before the first ace, is

$(51-n)(50-n)(49-n)/(51 \,.\, 50 \,.\, 49 \,.\, 13)$.

If the cards are continued to be dealt until a second ace appears, show that the chance that exactly r cards are dealt in all before the second ace, is

$r(51-r)(50-r)/(50 \,.\, 49 \,.\, 17 \,.\, 13)$.

Solution. (i) Exactly n cards are dealt before the first ace means that $(n+1)$th deal is the first ace. Thus n cards are extracted from the non-ace cards 48 (= 52 – 4) and $(n+1)$th card is extracted from $52-n$ cards. Define the events

$A = \{n$ cards out of 48 non-ace cards$\}$; $B = \{(n+1)$th card from $52-n$ cards$\}$.

$$P(AB) = P(A)\,P(B \mid A) = \left\{\binom{48}{n}\bigg/\binom{52}{n}\right\}\frac{4}{52-n}.$$

$$= \frac{48}{52!}\frac{(52-n)!}{(48-n)!}\frac{4}{52-n} = \frac{(51-n)(50-n)(49-n)}{51.50.49.13}.$$

(ii) Here $(r+1)$th card is to be the second ace; thus there is one ace in the first r cards. Define

$C = \{(r-1)$ cards out of 48 cards and one ace$\}$, $D = \{(r+1)$th card second ace$\}$

$$\therefore \qquad P(CD) = P(C)\,P(D \mid C) = \left\{\binom{4}{1}\binom{48}{r-1}\bigg/\binom{52}{r}\right\}.\frac{3}{52-r}$$

$$= 12.\frac{48!}{52!}\frac{(52-r)!}{(49-r)!}\frac{r!}{(r-1)!} = \frac{r(51-r)(50-r)}{50.49.17.13}.$$

***Example* 8.** Show that $P(A) = P(B)$ if $P(\bar{A}B) = P(A\bar{B})$.

Solution. From point-set algebra (draw a Figure) $A = A\bar{B} \uplus AB,\ B = \bar{A}B \uplus AB$.

$$\therefore \qquad P(A) = P(A\bar{B}) + P(AB); \quad P(B) = P(\bar{A}B) + P(AB),$$

Hence $\quad P(A) - P(B) = P(A\bar{B}) - P(\bar{A}B) \qquad \ldots(1)$

From (1), $\quad P(A) = P(B) \Leftrightarrow P(A\bar{B}) = P(\bar{A}B).$

Problems with Solutions Provided at the End of the Text

1*. Let $P(A) = p$, $P(A \mid B) = q$, $P(B \mid A) = r$. Find the relations between the numbers p, q, r for the following cases :

(a) Events A and B are mutually exclusive (m.e.).

(b) A and B are m.e. and collectively exhaustive.

(c) A is a subevent of B ; B is a subevent of A.

(d) $\bar{A}$ and $\bar{B}$ are m.e.

2*. If A, B, C are possible events, prove or disprove :
If $P(A) > P(B)$, then $P(A \mid C) > P(B \mid C)$.

3*. Show that if A, B, C are three events such that $ABC \neq \varnothing$ and $P(C \mid AB) = P(C \mid B)$, then $P(A \mid BC) = P(A \mid B)$.

4*. Suppose $B_1 \cap B_2 \neq \varnothing$, and $A_1 \cup A_2 \subset B_1B_2$. Show that
$P(A_1 \mid B_1)\, P(A_2 \mid B_2) = P(A_1 \mid B_2)\, P(A_2 \mid B_1)$.

5*. Let $B_1, B_2, \ldots, B_n$ be mutually disjoint and let $B = \uplus_j B_j$, $(1 \leq j \leq n)$. Suppose $P(B_j) > 0$ and $P(A \mid B_j) = p$ for $j = 1, 2, \ldots, n$. Show that $P(A \mid B) = p$.

6*. A coin is tossed until a head appears, or until it has been tossed three times. Given that the head does not appear on the first toss, find the probability that the coin is tossed three times.

7*. Two cards are drawn from a 52-card deck (the first is not replaced). Find the probability that

(a) If the first card is a jack, the 2nd is also a Jack. (b) Both cards will be a Jack.

(c) If the first card is a Jack and the 2nd card is a 10.

8*. A bag contains 15 items, of which 4 are defective. The items are selected at random, one by one and examined. The ones examined are not put back. What is the chance that the 10th one examined is the last defective ?

9*. An urn contains ten balls of which three are black and 7 are white. The following game is played : At each trial a ball is selected at random, its colour noted, and it is replaced along with two additional balls of the same colour. What is the probability that a black ball is selected in each of the first three trials ?

10*. The face cards are removed from a full pack. Out of the remaining 40 cards, 4 are drawn at random. Find the probability that :

(a) They belong to different suits.

(b) They belong to different suits and different denominations.

11*. Compare the probability of at least one six in 4 tosses of a fair die with the probability of at least one double-six in 24 tosses of two fair dice. (Chevalier de Mere Problem).

12*. A bag contains 50 balls of which three are black. The balls are drawn in succession (without replacement) from the bag. Find the chance that the first black ball is drawn at the rth draw. How is the probability changed if each ball is replaced before the next draw ?

13*. What is the chance that at least one of the players in a bridge game gets a complete suit of cards ?

14*. Define equivalent events. Prove that $P(ABC) = P(A \cup B \cup C)$, iff the events A, B, C are equivalent.

Exercise 1(d)

1. (a) Given $P(A) = 1/3$, $P(B) = 3/4$, $P(A \cup B) = 11/12$, show that $P(A \mid B)$ equals 2/9 and $P(B \mid A)$ equals 1/2.

   (b) Given $P(A) = 1/4$, $P(A \mid B) = 1/3$, $P(B \mid A) = 1/2$, find $P(A \mid \bar{B})$.

   (c) Given $P(A) = 1/2$, $P(B \cup A) = 3/5$, find $P(B)$ if $P(A \mid B) = 2/5$.

   (d) If $P(A) = P(B) = 1/2$, show that $P(A \mid B) \geq 1/2$.

   (e) If $P(B) = P(A \mid B) = P(C \mid AB) = 1/2$, show that $P(ABC) = 1/8$.

   (f) An experiment resulted in six outcomes $e_1, e_2, \ldots, e_6$ with respective probabilities 0.1, 0.2, 0.3, 0.2, 0.1, 0.1. If $A = \{e_1, e_3, e_4\}$, $B = \{e_2, e_4, e_6\}$ and $C = \{e_2, e_4, e_5\}$ show that (i) $P(A \cup B \cup C) = 1$, (ii) $P(A \mid B) = 2/5$, (iii) $P(\bar{B} \mid AC) = 0$.

2. (a) Show that $P(A \mid BC)\, P(B \mid C) = P(AB \mid C)$.

   (b) If B and C are subevents of an event A, show that $P(B \mid A)/P(C \mid A) = P(B)/P(C)$.

   (c) Suppose that $B \subseteq C$ with $P(B) > 0$. Show that

   (i) $P(A \mid B) = P(AB \mid C)/P(B \mid C)$, (ii) If $P(A \mid C) = 1$, then $P(A \mid B) = 1$.

   (d) Show that $P(A) + P(B) > 1 \Rightarrow P(B \mid A) \geq 1 - [P(\bar{B})/P(A)]$.

3. For the possible events A and B, establish the following :

   (a) $P(A \mid B) = P(B \mid A) \Leftrightarrow P(A) = P(B)$. (b) $P(A \mid B) = P(A) \Leftrightarrow P(B \mid A) = P(B)$.

   (c) $P(A \mid B) > P(A) \Leftrightarrow P(B \mid A) > P(B)$. (d) $P(A \mid B) < P(A) \Leftrightarrow P(B \mid A) < P(A)$.(B)

   Note. We say : B is *favourable* to A if $P(A \mid B) > P(A)$ and B is *unfavourable* to A if $P(A \mid B) < P(A)$. Thus, B carries *positive information* about A if $P(A \mid B) > P(A)$ and B carries negative information about A if $P(A \mid B) < P(A)$.

   Part (c) says : B is (unfavourable) to A iff A is (unfavourable) to B.

4. (a) Prove or disprove : If $P(A \mid B) = 1$, then $P(ABC) = P(BC)$ for any event C.

   (b) For independent events A and B, let $P(A\bar{B}) = x$, $P(\bar{A}B) = y$, where $0 \leq x, y \leq x + y \leq 1$. Find $p = P(AB)$. State additional restrictions on x, y if any.

5. (a) If $A \cap B = \varnothing$, then

   (i) $P(A \mid \bar{B}) = P(A)/[1 - P(B)]$, $P(B) \neq 1$

   (ii) $P(A \mid A \uplus B) = P(A)/[P(A) + P(B)]$, $P(A \uplus B) \neq 0$.

   (b) Show that $P(A \mid B) > 1 - P(\bar{A})/P(B)$.

   Show that the estimate of $P(A \mid B)$ from above and from below given that $P(A) = p$, $P(B) = 1 - \varepsilon$, where ε is small is $(p - \varepsilon)/)(1 - \varepsilon) \leq P(A \mid B) \leq p/(1 - \varepsilon)$.

   (c) For n events $A_1, A_2, \ldots, A_n$ with conditioning event B, $P(B) > 0$, show that

$$P\left(\bigcap_{i=1}^{n} A_i \mid B\right) \geq \sum_{i=1}^{n} P(A_i \mid B) - (n - 1).$$

6. Suppose B_j, $j = 1, 2, 3$ are m.e. events with $P(B_j) = 1/3$. If $P(A \mid B_j) = j/6$, then $P(A) = 1/3$.

7. An electronic assembly consists of two sub-systems, say A and B. From previous testing procedures, the following probabilities are assumed to be known :
$P(A \text{ fails}) = 0.20$, $P(A \text{ and } B \text{ fails}) = 0.15$, $P(B \text{ fails alone}) = 0.15$. Show that :
(i) $P(A \text{ fails} \mid B \text{ fails}) = 0.50$ (ii) $P(A \text{ fails alone}) = 0.05$.

8. Let A, B, C be three events such that :
$P(A) = 0.3$, $P(B) = 0.4$, $P(C) = 0.5$, $P(AB') = 0.2$, $P(BC) = 0.3$, $P(A'B'C') = 0.3$, $P(AB \mid C') = 0.1$.
Find $P(B' \mid C')$, $P(A \mid B)$ and $P(A \mid C')$. **[Ans.** 4/5, 1/4, 3/10]

9. The probability that a man purchases brand A or B of cigarette is 0.5 and 0.3 respectively. He purchased one of the two brands on a particular day. Show that the probability that he purchased A is 5/8.

10. In a class there are 13 girls and 14 boys. Two students are selected one by one without replacement for the recitation of a lesson. Show that the probability of selecting a girl on the second occasion given that
(i) a boy was selected on the first occasion is 7/27.
(ii) a girl was selected on the first occasion is 2/9.

11. A bag contains 6 balls of different colours and a ball is drawn from it. A speaks truth twice out of 4 times and B speaks truth 7 times out of 10 times. If both A and B say that a red ball was drawn, show that the probability of their joint statement being true is 35/38.

12. Urn A contains 3 black and 2 red balls and urn B contains 4 black and 6 red balls. Urn A is selected with probability p and B is selected with probability q $(= 1 - p)$. Find the value of p that makes the probability of getting a black ball the same as if a single draw is made from urn C that contains 7 black and 8 red balls.

13. (a) A bag contains 3 white and 3 black balls. They are drawn successively (i) without replacement, (ii) with replacement. Show that the chance that the colours alternate are 1/10 and 1/32.

(b) A bag contains 10 white and 6 black balls. 4 balls are successively drawn out and not replaced. Show that the probability that they are alternately of different colours is 45/364.

14. A person takes 4 tests in succession. The probability of his passing first test is p, that of his passing each succeeding test is p or $p/2$ according as he passes or fails the preceding one. He qualifies provided he passes at least 3 tests. Show that his chance of qualifying is $p^3 + \frac{3}{2} p^3 q$.

15. (a) A student is to appear for two tests in which his respective chances of winning are 0.5 and 0.7 and losing both the tests is 0.2. Show that the probability that the student will win the test-2 when he has already won test-1 is 0.80.

(b) A and B are two very weak students of mathematics; their chances of solving a problem correctly are 1/8 and 1/12 respectively. To a given problem, they obtain the same answer. If $P\{\text{common mistake}\} = 1/1001$, show that the chance is 13/14 that their answer is correct.

16. The sample space consists of the integers from 1 to $2n$, which are assigned probabilities proportional to their logarithms. Find the probabilities and show that the conditional probability of the integer 2, given that an even integer occurs, is $\log 2/[n \log 2 + \log n!]$.

17. A man forgets the last digit of a telephone number, and dials the last digit at random. Show that the probability of calling no more than three wrong numbers is 0.3.

18. (a) An urn contains 2 black, 3 white and 4 red balls. We draw one ball at random and put it aside. Then we draw the next ball and so on. Show that the probability of drawing at first two black balls, then three white ones and finally the red ones is 1/1260.

(b) A box contains w white and b black balls. When two balls are randomly extracted, the probability that both are white is given to be 1/2. Show that min $w = 3$ and that if b is even, min $w = 15$. Find also the min $(w + b)$.

19. A bowl contains 10 chips. Four of the chips are red, 5 are white and 1 is blue. If three chips are taken at random and without replacement, prove that the chance that there is 1 chip of each colour relative to the hypothesis that there is exactly 1 red chip among the three is 1/3.

20. A drawer contains 8 pairs of socks. If 6 socks are taken at random and without replacement, show that the chance that there is at least one matching pair among these 6 socks is 111/143.

21. A lot of 100 items undergoes a selective inspection. The entire lot is rejected if there is at least one defective item in five items checked. Show that the probability that the given lot will be rejected if it contains 5% defective items is 0.23.

22. A loaded die is tossed once. If x is the result of the toss, then $P(X = k) = p_k$, $1 \le k \le 6$. If $X = k$, a coin is tossed k times. Find $P\{X = \text{odd} \mid \text{at least one head occurs}\}$.

23. A supplier of a certain testing device claims that his device has high reliability inasmuch as $P(A \mid B) = P(\bar{A} \mid \bar{B}) = 0.95$, where $A = \{\text{device indicates component is faulty}\}$ and $B = \{\text{component is faulty}\}$. Device is to be used for locating faulty components in a large batch of components of which 5% are faulty. Show that $P(B \mid A) = 0.5$.

If we want $P(B \mid A) = 0.9$ and let $P(A \mid B) = P(\bar{A} \mid \bar{B}) = p$, how large does p have to be.

[$p = 0.9942$]

1-70. De-Moivre's Problem on Dice Points

Given n dice, each with f faces marked from 1 through f. These are thrown at random. The chance that the sum of the numbers exhibited is 's' is given by

$$P_s = \text{Coefficient of } x^s \text{ in } (x^1 + x^2 + \ldots + x^f)^n/(f)^n. \qquad \ldots(1)$$

Proof. Since any one of the f faces may be exposed on any one of the n dice, the total number of points constituting the sample space is f^n [Product Rule].

Since the coefficent of x^s in the multinomial expansion of $(x + x^2 + \ldots + x^{(f)})^n$ arises out of the different ways in which n of the indices 1, 2, ..., f can be taken so as to yield the sum S, it follows that the number of ways favourable to get sum S is the coefficient of x^s in the said multinomial. Hence, the probability sought is that given by (1).

Cor. P_s = Coefficient of x^{s-n} in $(1 + x + \ldots + x^{f-1})^n$

= Coefficient of x^{s-n} in $(1 - x^f)^n (1 - x)^{-n} / f^n$

$$= \text{Coefficient of } x^{S-n} \text{ in } \frac{(1-x^f)^n}{f^n} \sum_{r=0}^{n} \frac{n(n+1)(n+2)\ldots(n+r-1)}{r!} x^r. \qquad \ldots(2)$$

Extension. Given n dice, with faces $f_1, f_2, \ldots, f_n$, then

$$P_S = \text{Coefficient } x^S \text{ in } \frac{(x^1 + x^2 + \ldots + x^{f_1})(x^1 + x^2 + \ldots + x^{f_2})\ldots(x^1 + x^2 + \ldots + x^{f_n})}{f_1 \cdot f_2 \ldots f_n}.$$

The proof is smilar to (1).

Remark. De-Moivre's Problem can be more naturally handled by Generating Functions. See §8-50.

Case of Two Cubical Dice :

$$p_k = \text{Coefficient of } x^k \text{ in } \frac{(x + x^2 + ... x^6)^2}{(6)^2}, \quad 2 \le k \le 12$$

$= \lambda$. Coefficient of x^{k-2} in $(1 + x + ... + x^5)^2$, $\qquad [\lambda = 1/36]$

$= \lambda$. Coefficient of x^{k-2} in $[(1 - x^6)/(1 - x)]^2 = (1 - x^6)^2 . (1 - x)^{-2}$

$= \lambda$. Coefficient of x^{k-2} in $(1 - 2x^6 + x^{12}) \{\Sigma(r + 1)\, x^r\}$, $0 \le r < \infty$

$$= \text{Coefficient } x^{k-2} \text{ in } \lambda\left\{\Sigma(r+1)\, x^r - 2\Sigma\,(r+1)\, x^{r+6} + \Sigma(r+1)\, x^{r+12}\right\}.$$

We now put $k = 2, 3, ..., 12$, and record the coefficient to get

$p_2 = \lambda$, $p_3 = 2\lambda$, $p_4 = 3\lambda$, $p_5 = 4\lambda$, $p_6 = 5\lambda$, $p_7 = 6\lambda$, $p_8 = 5\lambda$

We can use $p_k = p_{7n-k} = p_{14-k}$ to recover

$p_8 = p_6$, $p_9 = p_5$, $p_{10} = p_4$, $p_{11} = p_3$, $p_{12} = p_2$.

All these results can be condensed into a single formula

$$\boxed{p_k = \frac{6 - |7 - k|}{36}, 2 \le k \le 12 \text{ or } p_k = p_{14-k} = \frac{k-1}{36}, k = 2, 3, ..., 7.}$$

Case of Three Cubical Dice :

p_k = Coefficient of x^k in $(x + x^2 + ... + x^6)^3/(6)^3$, $3 \le k \le 18$

$= \lambda$. Coefficient of x^{k-3} in $(1 + x + ... + x^5)^3$, $[\lambda = 1/216]$

$= \lambda$. Coefficient of x^{k-3} in $[(1 - x^6)/(1 - x)]^3 = (1 - x^6)^3\, (1 - x)^{-3}$

$= \lambda$. Coefficient x^{k-3} in $[1 - 3x^6 + 3x^{12} - x^{18})\, [\Sigma\{(r + 1)\,(r + 2)/2]\, x^r\}$, $0 \le r \le \infty$

$=$ Coefficient of x^{k-3} in $\{\Sigma\, \phi(r)\, x^r - 3\Sigma\, \phi(r)\, x^{r+6} + 3\, \Sigma\, \phi(r)\, x^{r+12} - \Sigma\, \phi(r)\, x^{r+18}\}$

where $\qquad \phi(r) = (r + 1)\,(r + 2)/2$. Note that

$\qquad \phi(0) = 1$, $\phi(1) = 3$, $\phi(2) = 6$, $\phi(3) = 10$, $\phi(4) = 15$, ...

We now put $k = 3, 4, ..., 10, ...$ and obtain

$p_3 = \lambda$, $p_4 = 3\lambda$, $p_5 = 6\lambda$, $p_6 = 10\lambda$, $p_7 = 15\lambda$, $p_8 = 21\lambda$, $p_9 = 25\lambda$, $p_{10} = 27\lambda$

Other values can also be obtained from complement rule.

$\qquad p_k = p_{7n-k} = p_{21-k}$.

Thus, $\qquad p_3 = p_{18}$, $p_4 = p_{17}$, $p_5 = p_{16}$, $p_6 = p_{15}$, $p_7 = p_{14}$, $p_8 = p_{13}$, $p_9 = p_{12}$, $p_{10} = p_{11}$

These results can be condensed to

$$\boxed{p_k = \frac{(k-1)(k-2)}{2 \times 216}, 3 \le k \le 8; \quad p_k = \frac{21(k-4) - (k^2 - 1)}{216}, 9 \le k \le 14.}$$

Illustrations. (i) For three dice :

$\qquad p_9 : p_{10} = 27\, \lambda : 25\, \lambda = 27 : 25 \qquad$ [*Galileo Problem*]

(ii) $P\{\text{Sum more than } 14\}$

$$p = p_{15} + p_{16} + p_{17} + p_{18} \quad \text{[Apply Complement Rule]}$$
$$= p_6 + p_5 + p_4 + p_3 = 20\lambda = 20/216 = 5/54.$$

(iii) A and B throw 3 dice. If A throws 8, what is B's chance of throwing a higher number ?

$$p = P\{\text{sum} > 8\} = 1 - P\{\text{Sum} \le 8\} = 1 - \{p_3 + p_4 + p_5 + p_6 + p_7 + p_8]$$
$$= 1 - (\lambda + 3\lambda + 6\lambda + 10\lambda + 15\lambda + 21\lambda) = 1 - 56\lambda$$
$$= 1 - \left(\frac{56}{216}\right) = 1 - \left(\frac{7}{27}\right) = \frac{20}{27}.$$

***Example* 1.** Find the chance of throwing more than 14 with three symmetrical dice.

Solution. Here $f = 6$, $n = 3$, $S > 14$; let $P(\text{Sum} = S) = P_S$, then

$$P_S = \text{Coefficient of } x^s \text{ in } (x + x^2 + \ldots + x^6)^3/(6)^3$$
$$= \text{Coefficient of } x^{s-3} \text{ in } (1 + x + \ldots + x^5)^3/216$$
$$\therefore \; 216\, P_S = \text{Coefficient of } x^{s-3} \text{ in } (1 - x^6)^3 (1 - x)^{-3}$$
$$= \text{Coefficient of } x^{s-3} \text{ in } (1 - 3x^6 + 3x^{12} - x^{18})\,(1 + 3x + \ldots + \tfrac{1}{2}\,(r+1)\,(r+2)\,x^r + \ldots.)$$

For the present situation, we require Coefficient of $x^{12}, x^{13}, x^{14}, x^{15}$, [for $S = 15, 16, 17, 18$].

Coefficient of $x^{12} = 91 - 84 + 3 = 10$, Coefficient of $x^{13} = 105 - 108 + 9 = 6$

Coefficient of $x^{14} = 120 - 135 + 18 = 3$, Coefficient of $x^{15} = 136 - 165 + 30 = 1$

$$\therefore \; P(S > 14) = p_{15} + p_{16} + p_{17} + p_{18} = (10 + 6 + 3 + 1)/216 = 5/54.$$

***Example* 2.** Compare the probability of a total of 9 with a total of 10 when three fair dice are rolled once.

Solution. Here $n = 3, f = 6, S_1 = 9, S_2 = 10$. Now

$$p_9 = P(S = 9) = \text{Coefficient of } x^9 \text{ in } (x + x^2 + \ldots + x^6)^3/(6)^3$$
$$= \text{Coefficient of } x^6 \text{ in } (1 - x^6)^3 (1 - x)^{-3}/216.$$
$$\therefore \; 216p = \text{Coefficient of } x^6 \text{ in } (1 - 3x^6 + \ldots)\,[1 + 3x + \ldots + (\tfrac{1}{2})(r+1)\,(r+2)x^r + \ldots] = 28 - 3 = 25.$$
$$\therefore \; P(S = 9) = 25/216$$
$$p_{10} = P(S = 10) = \text{Coefficient of } x^{10-3} \text{ in } (1 - x^6)^3 (1 - x)^{-3}/216. \quad \text{[as above]}$$
$$\therefore \; 216p_{10} = \text{Coefficient of } x^7 \text{ in } (1 - 3x^6 + \ldots)\,[1 + 3x + \ldots + (\tfrac{1}{2})(r+1)\,(r+2)\,x^r + \ldots]$$
$$= 36 - 9 = 27.$$

Thus $P(S = 10) = 27/216$. It follows that $P(S = 9) : P(S = 10) = 25 : 27$.

Exercise 1(e)

1. **Complement Rule.** Show that the probability of throwing a total of S with n ordinary six sided dice is the same as the probability of throwing a total of $(7n - S)$.
2. Prove that the chances of throwing 10 in one throw with (a) 3 dice, (b) 4 dice are 1/8 and 5/81.
3. Three fair dice are thrown once. If no two show the same face, prove that P {sum of faces = 7} = 1/36; P {one is an ace}.
4. Counters marked 1, 2, 3 are placed in a bag, and one is drawn and replaced. The process is repeated thrice. Show that the chance of obtaining a total of 6 is 7/27.

5. Four tickets marked 00, 10, 01, 11 respectively are placed in a bag. A ticket is drawn at random 5 times, being replaced each time. Prove that the chance of obtaining a total of 23 is 25/256.
6. Out of 10 tickets, 5 are blank and the rest are marked 1, 2, 3, 4, 5. Show that the chances of drawing a total of 10 in three trials when (a) tickets are replaced after every trial, (b) tickets are not replaced are (i) 33/1000, (ii) 1/60.
7. Nine cards are drawn from a set of cards. Each is marked with one of the numbers 1, 0, –1 and it is equally likely that any of the three numbers will be drawn. Prove that the chance of obtaining a total zero is $3139/(3)^9$.

1-71. *Probability and Number-arithmetic*

***Example* 1.** *Chebyshev's Problem.* What is the chance that two numbers, chosen at random, will be prime to each other ?

Solution. When an arbitrary number n is divided by any prime number p, the remainders are 0, 1, 2, ... $p - 1$. So if A is the event that n is divisible by p, $P(A) = (1/p)$ since only one case, viz. zero is favourable to A. Consequently, if two numbers x and y are both divisible by p, then

$$P(AB) = P(A)\,P(B) = (1/p)\,(1/p) = (1/p)^2.$$

It follows that, both x and y are not divisible by p has the probability $1 - (1/p^2)$. So, for every two relatively prime numbers, the required probability is

$$P = \prod_{p=0}^{\infty}\left[1-\frac{1}{p^2}\right]=\frac{6}{\pi^2} \qquad \text{[From Trigonometry]}$$

***Example* 2.** A five-figure number is formed by the digits 0, 1, 2, 3, 4 (without repetitions). Find the chance that the number formed is divisible by 4.

Solution. We can arrange 5 digits in 5 ! ways, but if the numbers start with 0, the effective arrangements reduce to $N = 5\,! - 4\,! = 96$, because there are 4 ! arrangements corresponding to the numbers 0 *abcd*.

Now the numbers ending in 04, 12, 20, 24, 32, 40 are obviously divisible by 4. Consequently, we find arrangements with these endings : In

xyz 04 : x, y, z can be arranged in 3 ! ways; *xyz* 40 and *xyz* 20 : x, y, z each can be arranged in 3! ways

xyz 12 : x, y, z can be arranged in $3\,! - 2\,! = 4$ ways. [$x = 0$, extreme-left, reduces the cases]

xyz 24 : x, y, z can be arranged in $3\,! - 2\,! = 4$ ways.

xyz 32 : x, y, z can be arranged in $3\,! - 2\,! = 4$ ways.

xyz 40 : x, y, z can be arranged in 3 ! ways.

Total number of favourable cases, by Rule of Sum, is 30. Hence $p = 30/96 = 5/16$.

***Example* 3.** Digits 1, 2, 3, 4, 5, 6, 7 are written down in any order, [each possible order being equally likely] to form a seven-digit number. Find the chance that this number is divisible by eight.

Solution. The total number of equally likely seven-digit numbers is $N = 7\,! = 5040$.

We find the number of favourable cases M included in N which are divisible by eight. Note that 1,000 and hence 10,000, 100,000 and 1,000,000 are all exactly divisible by

eight, and that 100, 10, 1 leave respective remainders 4, 2, 1 on division by 8. It follows that when the seven-digit number *abcdefg* is divided by 8 there will be a remainder of the form $R = 4e + 2f + g$. Thus we need calculate the *different possible values* of e, f, g which make R an exact multiple of 8. Triplet (x, y, z) means $g = x$, $f = y$, $e = z$. The sets of value rendering R an exact multiple of 8 are :

(2, 1, 3), (2, 1, 5), (2, 1, 7), (2, 3, 4), (2, 3, 6), (2, 5, 1), (2, 5, 3), (2, 5, 7), (2, 7, 4), (2, 7, 6); (4, 2, 6), (4, 6, 2); (6, 1, 2), (6, 1, 4), (6, 3, 1), (6, 3, 5), (6, 3, 7), (6, 5, 2), (6, 5, 4), (6, 7, 1), (6, 7, 3), (6, 7, 5).

These are 22 triplets indicating 22 different favourable sets of values of e, f, g. For each triplet, there are 4 ! different orders of the remaining four digits a, b, c, d. Thus $M = 22 \times 4\,!$ (favourable seven-digit numbers). Consequently

$$p = \frac{M}{N} = \frac{22 \times 4!}{7!} = \frac{11}{105}.$$

Exercise 1(f)

1. Show that the probability that a positive integer selected at random is relatively prime to 6 is 1/3. And that at least one of two integers selected at random is relatively prime to 6 is 5/9.
2. (a) Find the probability P_N that a natural number chosen at random from the set $\{1, 2, 3, \ldots, N\}$ is divisible by a fixed natural number k. Also find $\lim P_N$, as $N \to \infty$.

   (b) Find the chance that an integer chosen from the first 100 natural numbers is divisible by at least 2 of the four primes : 2, 3, 5, 7.
3. If 4 whole numbers taken at random are multiplied together, show that the chance that the last digit in the product is 1, 3, 7 or 9 is 16/625.
4. Four positive integers are chosen at random. Prove that the chance of their having a common factor is $1 - (90/\pi^4)$.
5. A number consists of 7 digits whose sum is 59. Prove that the chance of its being divisible by 11 is 4/21.
6. One card is selected at random from 100 cards numbered 00, 01, ..., 98, 99. Suppose X and Y are the sum and product respectively of the digits on the selected card. Find $P(X = k \mid Y = 0)$. [$p = 1/19, k = 0\,; p = 2/19, k = 1, 2, \ldots, 9$].
7. The numbers 2, 4, 6, 7, 8, 11, 12, 13 and 17 are written respectively, on 10 indistinguishable cards. Two cards are selected at random from ten. Show that the probability, that the fraction formed with them is reducible, is 4/61.
8. If a 3-digit number (000 to 999) is chosen at random, find the chance that exactly 1 digit will be greater than k.
9. Show that the chance that the last two digits of the cube of a 4-digit random integer will be two ones is 0.01. [We assume that any four-digit integer has the same chance to be chosen].
10. Out of $6n$ tickets numbered $0, 1, 2, \ldots, 6n - 1$, three are drawn at random. Show that the chance that their sum is $6n$, is given by $3n/(6n-1)(6n-2)$.
11. An integer expressed in the decimal system requires five digits. Stating reasonable assumptions, find the chance of guessing the number.
12. Let $\phi(n)$ denote the number of positive integers which are less than n and which are primes relative to n. Prove that $\phi(n) = n\prod(1 - p^{-1})$ where product extends over all prime divisors p of n. [$\phi(n)$ is Euler's function].

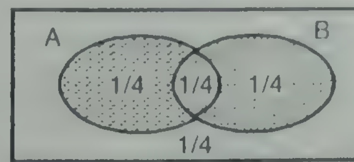
1-80. Miscellaneous Worked-out Problems

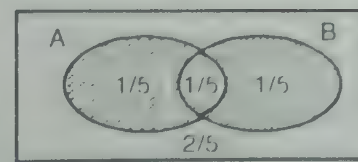
***Example* 1.** Show by examples that $P(B \mid A) + P(B \mid \bar{A})$

(i) may be equal to 1,

(ii) may not be equal to 1.

However $P(A \mid B) + P(\bar{A} \mid B) = 1$ is true.





Solution. We assign probabilities as under

(i) $P(A\bar{B}) = P(AB) = P(\bar{A}B) = P(\bar{A}\bar{B}) = \frac{1}{4}$

Then
$$p = P(B \mid A) + P(B \mid \bar{A}) = \frac{P(AB)}{P(A)} + \frac{P(\bar{A}B)}{1-P(A)} = \frac{\frac{1}{4}}{\frac{1}{2}} + \frac{\frac{1}{4}}{\frac{1}{2}} = 1.$$

(ii)
$$P(A\bar{B}) = P(AB) = P(\bar{A}\bar{B}) = \frac{1}{5}, \qquad P(\bar{A}\bar{B}) = \frac{2}{5}$$

$$p = P(B \mid A) + P(B \mid \bar{A}) = \frac{P(AB)}{P(A)} + \frac{P(\bar{A}B)}{1-P(A)} = \frac{1/5}{2/5} + \frac{1/5}{3/5} = \frac{5}{6} \neq 1.$$

(iii)
$$P(A \mid B) + P(\bar{A} \mid B) = \frac{P(AB)}{P(B)} + \frac{P(\bar{A}B)}{P(B)} = \frac{P(AB) + P(\bar{A}B)}{P(B)} = \frac{P(B)}{P(B)} = 1.$$

***Example* 2.** A die is thrown as long as necessary for a 6 to turn up. Given that the 6 does not turn up at the first throw, find the chance that the more than four throw will be necessary.

Solution. That more than four throws are needed means that in the first 4 throws there is no 6 occurring. Let N_i denote the event that 6 does not occur on throw i, $(i = 1, 2, 3, \ldots)$ Trivially, $P(N_1) = 1$. The desired probability is thus

$$P\{N_1 \cdot N_2 \cdot N_3 \cdot N_4 \mid N_1\} = P(N_1)\, P(N_2 \mid N_1)\, P(N_3 \mid N_1 N_2)\, P(N_4 \mid N_1 \cdot N_2 \cdot N_3)$$
$$= (5/6) \cdot (5/6) \cdot (5/6) = 125/216.$$

Note that $N = N_1 \cdot N_2 \cdot N_3 \cdot N_4$ and N_1 are not independent, for $P(N \mid N_1) \neq P(N)$.

***Example* 3.** Out of n products, each with a distinct coupon, one has collected a set of r coupons, each of which is independently a type j coupon (labelled C_j) with probability p_j, where $p_1 + p_2 + \ldots + p_n = 1$. Find the probability that this set S contains coupon C_k given that S contains coupon C_j $(j \neq k)$.

Solution. Write $C_j = \{$Set S has coupon $C_j\}$. Then we need

$$P\{C_k \mid C_j\} = P\{C_j C_k\}/P\{C_j\}. \tag{1}$$

Now
$$P\{C_k\} = 1 - P\{\bar{C}_k\} = 1 - P\{\eta_0\ C_k \text{ coupon in } S\}$$
$$= 1 - (q_k^r) \quad [p_k + q_k = 1]$$
$$P\{C_j C_k\} = 1 - P\{\bar{C}_j \cup \bar{C}_k\} \qquad [P(A) = 1 - P(\bar{A})]$$
$$= 1 - \{P(\bar{C}_j) + P(\bar{C}_k) - P(\bar{C}_j \bar{C}_k)\}$$

As $$P\{\bar{C}_j\,\bar{C}_k\} = 1 - \{q_j^r + q_k^r - [1 - p_j - p_k]^r\}$$

Substituting into (1) yields

$$P\{C_k \mid C_j\} = \{1 - q_j^k - q_k^r + (1 - p_j - p_k)^n\} / \{1 - q_j^r\}.$$

***Example* 4.** A coin is tossed $(m + n)$ times, $m > n$. Show that the probability of at least m consecutive heads is $(n + 2)/2^{m+1}$.

Solution. Let H and T denote the occurrence of head and tail respectively. Let Y denote the occurrence of head or tail so that

$$P(H) = \tfrac{1}{2} = P(T);\ P(Y) = P(H \uplus T) = P(H) + P(T) = 1.$$

Let $Z_m = H \cap H \cap \ldots \cap H$ (occurrence of m consecutive heads), Y_j denoting head or tail on the jth throw. The possible sequences of interest are

$S_1 = Z_m\, Y\, Y \ldots Y,\qquad S_2 = TZ_m\, Y\, Y\, Y \ldots Y,\qquad S_3 = YTZ_m\, Y \ldots Y$

$S_r = Y.Y \ldots TZ_m\, Y\, Y \ldots Y, \ldots$ $[(r-1)$ terms before $Z_m]$, $S_{n+1} = Y \ldots Y.\, TZ_m$ (n terms before Z_m)

$$P(S_1) = P(Z_m) \,.\, [P(Y)]^n = (1/2)^m$$

$$P(S_2) = P(T)\,.\,P(Z_m)\,[P(Y)]^{n-1} = (1/2)\,(1/2)^m\,.\,(1) = (1/2)^{m+1}$$

$$P(S_r) = [P(Y)]^{n-1}\,.\,P(T)\,.\,P((Z_m).\,\{P(Y)]^{n-r} = 1\,.\,(1/2)\,.\,(1/2)^m\,.\,1 = (1/2)^{m+1}$$

$$P(S_{n+1}) = [P(Y)]^{n-1}\,.\,P(T)\,.\,P(Z_m) = 1.\,(1/2)\,(1/2)^m = (1/2)^{m+1}$$

$$p = P\{S_1 \uplus S_2 \uplus \ldots \uplus S_{n+1}\} = (1/2)^m + n(1/2)^{m+1} = (n+2)/(2)^{m+1}.$$

***Example* 5. Match-Mismatch Problem.** 'n' different objects 1, 2, ..., n are thrown at random in n places marked 1, 2, 3, ..., n. Find the probability that none of the objects occupies the place corresponding to its number. Also find P\{Exactly r objects occupy their matching places\}.

Solution. Write M_i = \{object i falls in slot i\}. That is M_i denotes the match occuring at slot i. Then probability of at least one match $P\{L_1\}$ is

$$P\{M_1 \cup M_2 \cup \ldots \cup M_n\} = \sum_i P(M_i) - \sum_{i<j} P(M_i\, M_j) + \sum_{i<j<r} P(M_i\, M_j\, M_k) - \ldots$$

$$+ (-1)^{r-1} \Sigma\,\Sigma \ldots \Sigma\, P(M_{i_1}\, M_{i_2} \ldots M_{i_r}) + \ldots + (-1)^{n-1} P(M_1\, M_2 \ldots M_n) \quad \ldots(1)$$

Now $$P\{M_i\} = \frac{1}{n} = \frac{(n-1)!}{n!} \qquad \text{[one favourable slot out of } n]$$

$$P\{M_i M_j\} = P\{M_i\}\, P\{M_j \mid M_i\} = \frac{1}{n}\cdot\frac{1}{n-1} = \frac{(n-2)!}{n!}$$

$$P\{M_i M_j M_k\} = P\{M_i\}\, P\{M_j \mid M_i\}\, P\{M_k \mid M_i M_j\} = \frac{1}{n}\cdot\frac{1}{n-1}\cdot\frac{1}{n-2} = \frac{(n-3)!}{n!}$$

In general : $$P\{M_{i_1}, M_{i_2}, \ldots, M_{i_r}\} = \frac{(n-r)!}{n!},\quad r = 1, 2, \ldots, n. \qquad \ldots(2)$$

Substituting into the Union Rule (1) and using symmetry of evaluations gives

$$P(L_1) = \binom{n}{1}\frac{(n-1)!}{n!} - \binom{n}{2}\frac{(n-2)!}{n!} + \ldots + (-1)^{r-1}\binom{n}{r}\frac{(n-r)!}{n!} + \ldots + (-1)^{n-1}\binom{n}{n}\frac{(n-n)!}{n!}$$

Using $\binom{n}{r} = \frac{n!}{r!(n-r)!}$, this simplifies to

$$P\{L_1\} = 1 - \frac{1}{2!} + \frac{1}{3!} - \ldots + \frac{(-1)^{k-1}}{k!} - \ldots + \frac{(-1)^{n-1}}{n!} = \sum_{k=1}^{n} \frac{(-1)^{k-1}}{k!}.$$

The probability of no matches is given by

$$P\{\overline{M}_1\ \overline{M}_2 \ldots \overline{M}_n\} = 1 - P\{L_1\} = 1 - \sum_{k=1}^{n} \frac{(-1)^{k-1}}{k!} = \sum_{k=0}^{n} \frac{(-1)^k}{k!}. \qquad \ldots(3)$$

Second Part. Probability of exactly r matches $= P(E_r)$.

We consider all possible arrangements of r successes in n trials, viz. $\binom{n}{r}$ and hence

$$P\{E_r\} = \binom{n}{r} P\left\{M_1\ M_2 \ldots M_r\ \overline{M}_{r+1}\ \overline{M}_{r+2} \ldots \overline{M}_r\right\}$$

$$= \binom{n}{r} P\{S_r\ F_{n-r}\}, \qquad [S_r = M_1\ M_2 \ldots M_r,\ F_{n-r} = \overline{M}_{r+1} \ldots \overline{M}_n]$$

$$= \binom{n}{r} P\{S_r\}\, P(F_{n-r} \mid S_r). \qquad \text{[By Product Rule]} \qquad \ldots(4)$$

Now $\quad P(S_r) = \frac{(n-r)!}{n!} \qquad$ [by Eq. (2)]

$$P\{F_{n-r} \mid S_r\} = P\{\text{All } (n-r) \text{ objects mis-match}\} = \sum_{k=0}^{n-r} \frac{(-1)^k}{k!} \qquad \text{[by (3)]}$$

Substituting these evaluations into (4) yeilds

$$P\{E_r\} = \frac{n!}{r!(n-r)!} \cdot \frac{(n-r)!}{n!} \sum_{k=c}^{n-r} \frac{(-1)^k}{k!}$$

$$\boxed{P\{E_r\} = \frac{1}{r!} \sum_{k=0}^{n-r} \frac{(-1)^k}{k!}} \qquad \ldots(5)$$

Example 6. Each coefficient in the equation $ax^2 + bx + c = 0$, is determined by throwing an ordinary die. Find the probability that the equation will have (i) real roots, (ii) complex roots.

Solution. Coefficient a, b, c are each uniformly distributed over $\{1, 2, 3, 4, 5, 6\}$; hence the exhaustive number of cases $N = 6 \times 6 \times 6 = 216$. [The three coefficients a, b, c are realy determined by throwing a die three times]. The roots of the given equation are real iff $4ac \le b^2$. Since max $b = 6$, min $b = 2$ (to satisfy the constraint), the favourable cases are as under :

b	$ac \le b^2/4$	*Outcomes* ω	*Frequency*
2	$ac \le 1$	(1, 1)	1
3	$ac \le 2$	(1, 2), (2, 1), (1, 2)	3
4	$ac \le 4$	(1, 1), ..., (1, 4), (2, 1), (2, 2), (3, 1), (4, 1)	8
5	$ac \le 6$	(1, 1), ..., (1, 6), (2, 1), (2, 2), (2, 3); (3, 1), (3, 2), (4, 1), (5, 1), (6, 1)	14
6	$ac \le 9$	(1, 1), ..., (1, 6); (2, 1), ..., (2, 4); (3, 1), (3, 2), (3, 3), (4, 1), (4, 2), (5, 1), (6, 1)	17

Thus total No. of favourable cases = 1 + 3 + 8 + 14 + 17 = 43. So $p = 43/216$.

The probability that the roots are complex is $q = 1 - p = 173/216$.

Problems with Solutions Provided at the End of the Text

1*. If $A \cap B = \varnothing$, show that $P(A) \leq P(\bar{B})$ and $P(B) \leq P(\bar{A})$.

2*. An urn contains n white balls numbered 1 through n, n black balls numbered 1 through n and n red balls numbered to 1 to n. Two balls are drawn at random without replacement. Find the chance that both balls are of the same colour or bear the same number.

3*. The Mathematician, Professor S.S. Khurana, travels from Iowa City to New Delhi with stop-overs in New York and London. At each stop his luggage is transferred from one plane to another. In each airport, including Iowa, chances are that with probability q his luggage is not placed in the right plane. Prof. Khurana finds that his luggage has not reached Delhi. Find the probabilities that the misplacement took place in Iowa, in the New York and in London, respectively.

1-90. *Comments on 'Surprise' and 'Uncertainty'*

Naively, the surprise one feels upon learning that an event A has happened depends on $P(A) = p$ (say), $0 < p \leq 1$ and thus we denote the Surprise function by $S(p)$. Obviously, $S(p) = 0$ if $p = 1$ (certainty) and our surprise decreases if p increases. Further, span of p maintains span of $S(p)$ and occurrences of additional events adds to additional surprises. These facts lead to the definition.

$$S(p) = -\ell n_a\, p \;(a > 1) \qquad \text{[Structure of } S(p)\text{]}$$

The *Uncertainty* or *Entropy*, denoted by H is defined by

$$H = E[S(p)] = E(-\ell n_a\, p).$$

Thus, for the probability ensemble (X, p) ; $X = x_1, \ldots, x_n$ and $p = p_1, \ldots p_n \ldots$ we define

$$H(X) = E[S(p)] = -\sum p_i\, \ell n_a\, p_i.$$

Illustration. Urn A contains 60 red and 40 blue balls and urn B contains 6 red and 4 blue balls. An urn is chosen at random and 2 balls withdrawn simultaneously.

(i) You win a prize if *either* ball is red.

(ii) You win a prize if either ball in blue.

Which urn, A or B, gives you the best chances of winning the prize in case (i) ? in case (ii) ?

Solution. We evaluate entropies H_A and H_B for both cases

(i) $$P\{2 \text{ red balls}\}_A = \binom{60}{2}\Big/\binom{100}{2} = \frac{177}{495} = 0.36$$

$$P(2 \text{ red balls})_B = \binom{6}{2}\Big/\binom{10}{2} = \frac{1}{3} = 0.33$$

$$H_A = -p\,\ell n_2\, p = -(0.36)\,\ell n_2\,(0.36) = 0.53062$$
$$H_B = -p\,\ell n_2\, p = -(0.33)\,\ell n_2\,(0.33) = 0.52762$$

$H_A > H_B \Rightarrow$ urn B is of greater certainty for red balls.

(ii) $P\{2 \text{ blue balls}\}_A = \binom{40}{2} \Big/ \binom{100}{2} = \frac{78}{495} = 0.76$

$P\{2 \text{ blue balls}\}_B = \binom{4}{2} \Big/ \binom{10}{2} = \frac{2}{15} = 0.13$

$$H_A = -(0.76)\ \ell n_2\ 10.76 = 0.42302$$

$$H_B = -(0.13)\ \ell n_2\ (0.13) = 0.38264$$

$\therefore\ H_A > H_B \Rightarrow$ urn B is of greater certainty for blue balls.

Conclusion. In either case, (i) or (ii), urn B yields the best chances of winning the prize.

Miscellaneous Exercises

1. With usual notation, show that
   (a) If $P(A) = P(B) = 1$, then $P(AB) = 1$.
   (b) If $P(A_i) = 1, i = 1, 2, \ldots,$ then $P(\bigcap_i A_i) = 1$.
   (c) If $P(A \mid C) \geq P(B \mid C)$ and $P(A \mid \overline{C}) \geq P(B \mid \overline{C})$, then $P(A) \geq P(B)$.
   (d) Prove : $P(A \mid B) = P(A \mid BC)\, P(C \mid B) + P(A \mid B\overline{C})\, (\overline{C} \mid B)$.
   (e) Under what conditions does the following equality hold :
   $P(A) = P(A \mid B) + P(A \mid \overline{B})$. $[A = \varnothing, B = \Omega; B = A, B = \overline{A}, B = \varnothing]$
2. (a) The probability that a student passes Mathematics is 2/3, that of passing Biology is 4/9. If the probability of passing at least one course is 4/5, show that the probability that he will pass both courses is 14/45.
   (b) Eight persons are arranged in a row, all possible arrangements being equiprobable. Find the chance that two persons will be next to each other. What happens if arrangement be in a ring ?
3. The chance of first event happening is the square of the chance of a second event, but the odds against the first are the cube of the odds against the second. Find the chance of each. [**Ans.** 1/9, 1/3]
4. Two unbiased dice, one green and other red, are tossed simultaneously. Let g and r denote the numbers on green and red die respectively. Find
   (a) $P\{g \geq r + 4\}$ (b) $P\{g \neq r\}$ (c) $P\{g = r\}$
   (d) $P\{g = r^2\}$, (e) $P\{g + r \neq 7\}$, (f) $P\{g + r = 10\}$
   (g) $P\{g = 3r\}$ (h) $P\{g - r = 1\}$. [**Ans.** $3p, 30p, 6p, 2p, 30p, 3p, 2p, 5p, p = 1/36$]
5. A box contains 20 fuses of which 5 are defective. Show that the probability that the two fuses are defective when they are selected at random and removed from this box, without replacing the first is 1/19.
6. Four persons are chosen at random from a group of 4 men, 3 women and 5 children. Prove that the probability that exactly two of them will be children is 16/27.
7. You need 5 eggs to make omelet for break-fast. You find 12 eggs in the refrigerator but do not realize that 3 of them are rotten. Show that the probability that of the five eggs you choose (a) exactly one is rotten, $p = 21/44$, (b) none is rotten, $p = 7/44$.

8. (a) A die is rolled 3 times. Prove that the probability that you get a larger number each time is 20/216.

(b) A pair of dice is tossed as long as necessary for a score of 7 to turn up. If 7 does not appear on first throw, find the chance that more than two throws will be necessary.

(c) Two tetrahedra are tossed repeatedly until a total of 5 appears on the down faces that P {More than two tosses are required} = 9/16.

9. A king is placed at random on any square of a standard chess-board, other than the edge-squares. A queen is placed at random on any square of the same chess-board other than the square occupied by the king. Prove that the probability that the king and the queen are on neighbouring (horizontal, vertical or diagonal) squares is 8/63.

10. A bridge player knows that his two opponets have exactly five hearts between the two of them. Each opponent has 13 cards. Prove that the probability that there is a three-two split on the hearts is $2\binom{5}{2}\binom{21}{10}\Big/\binom{26}{13}$.

11. From 25 tickets marked with the first 25 numerals, one is drawn at random. Find the chance that it is a multiple of : (a) 6, (b) 3 or 7, (c) 5 or 7, (d) 3 and 6. [**Ans.** 0.16, 0.40, 0.32, 0.32]

12. Three distinct integers are chosen at random from the first 20 positive integers. Show that : (a) P {their sum is even} = 1/2, (b) P {their product is even} = 17/91.

13. (a) Four people are dealt five cards each. The first player gets four hearts and one club. He discards the club, and takes in exchange the top card from the remaining 32 cards. Show that the chance that he now has 5 hearts is 9/47.

(b) A trainee gunman fires n shells at a target. The chance of the kth shell hitting the target is p_k, $1 \le k \le n$. Find the chance that at least two shells will hit the target.

14. A bowl contains 16 balls of which 6 are red, 7 are white and 3 are blue. If 4 balls are taken at random and without replacement, show that :

(a) P{Each of the four balls is red} = 15/1820

(b) P{None of the balls is red} = 210/1820

(c) P{There is at least one ball of each colour} = 819/1820.

15. An urn contains five red and three white balls. The red balls are numbered 0, 1, 2, 3, 4 respectively and the white balls are numbered 0, 1, 2 respectively. Two balls are withdrawn without replacement. Prove that the chance that these balls have either the same number or the same colour is 4/7.

16. Each of the n urns contains three chips numbered 1, 2, 3. A chip is chosen at random from each urn yielding a set S of integers. Prove that the probability that this set contains each of the numbers 1, 2, 3 at least once is $1 - 3(2/3)^n + 3(1/3)^n$.

17. (a) Fifty-two persons stand in a line, each to draw without replacement, a card from an ordinary deck. The Queen of Spades is the prize card. What place in the line is the best ?

[**Ans.** All positions in the line are equally likely]

(b) If n people are seated around a round table, what is the chance that two named individuals will be next to each other ? [**Ans.** $2/(n-1)$]

18. In a T maze, a rat is given a choice of going to the left (L) and getting food or going to the right (R) and getting some shock. Before any conditioning (in trial No. 1) the rats are equally likely to go to L or to R. After getting food on a particular trial, the probability of going to L and to R become $p_1 > 1/2$ and q_1 on the following trial; but in case of a shock, these

probabilities are p_2 and q_2. Find the probability that rat will turn to L on trial number 2 and also on trial number 3. $(p_2 > p_1)$. Also find $P\{$rat turns to L on trial 3 | rat turned to L on trial 1$\}$.

19. A bag contains 6 white and 9 black balls and four balls are drawn at a time. Show that the probability p for the first draw to give 4 white balls and the second to give 4 black balls when :

(a) The balls are replaced before the 2nd draw $p = 6/5915$.

(b) The balls are not replaced before the 2nd draw $p = 3/715$.

20. An elevator starts with 7 passengers and stops at 10 floors. Show that

(a) $P\{$No two passengers leave at the same floor$\} = 109 \ldots 54/(10)^7$

(b) $P\{$All the passengers are discharged on the 5th floor$\} = (10)^{-7}$

(c) $P\{$All of the passengers are discharged on the same unspecified floor$\} = (10)^{-6}$

(d) $P\{$Exactly, three of them are discharged on the 2nd floor$\} = \binom{7}{3}(7)^4 \Big/ (10)^7$].

21. A bag contains 10 balls either black or white, but it is not known how many of each. A ball is drawn at random and it is found to be white. Assuming that all combinations of white and black balls are equally likely, show that the chance that the bag contains at least 5 white balls originally is 0.82.

22. A consignment of 15 record players contains 4 defectives. The record-players are selected at random, one by one, and examined. The ones examined are not put back. Prove that the probability that the 9th one examined is the last defective is 8/195.

23. An urn contains 5 balls. Two balls are drawn and found to be white. What is the probability of all the balls being white ?

24. From a basket containing 3 bad and 7 good oranges, two are selected at random. If the second selected is given to be bad, show that the chance, that the first selected is bad is 2/9.

25. Two ordinary six-sided dice are tossed. What is the probability that both the dice show (a) the number 4, (b) the same number. Given that the sum of the two numbers shown is 8, find the conditional probability that the number noted on the first die is as large as the number noted on the second die. **[Ans. 1/36, 6/36, 1/5]**

26. The probability that a person stopping at a Petrol Pump will ask to have its tyres checked is 0.12, the probability that he will ask to have its oil checked is 0.29 and the probability that he will ask to have them both checked is 0.07.

(a) What is the probability that a person stopping at this Petrol pump will have :

(i) Either his tyres or his oil checked ?

(ii) Either his tyres or his oil checked but not both ?

(b) What is the probability that a person who has his tyres (oil) checked will also have his oil (tyres) checked ? **[Ans. (a) 0.34, 0.41; (b) 7/29, 7/12]**

27. A hand of 13 cards is to be dealt at random and without replacement from an ordinary pack of playing cards. Find the conditional probability that there are at least 3 kings in the hand given that the hand contains at least 2 kings.

$$\left[\textbf{Ans. } p = \left\{\left[p = \binom{4}{3}\Big/\binom{48}{10} + \binom{4}{4}\binom{48}{9}\right]\Big/\left[\binom{4}{2}\binom{48}{11} + \binom{4}{3}\binom{48}{10} + \binom{4}{4}\binom{48}{9}\right]\right\}\right]$$

28. Four different objects 1, 2, 3, 4 are distributed at random on four places marked 1, 2, 3, 4. Prove that the probability that none of the objects occupies the place corresponding to its number is 3/8.

29. One term of the expansion of a determinant of the nth order is chosen at random. Show that the probability P_n that it does not contain element of the principal diagonal is

$$P_n = 1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \ldots + (-1)^n \frac{1}{n!} \qquad \left[\lim_{n\to\infty} P_n = e^{-1}\right]$$

[**Hint.** The number of the terms in the expansion of a Det. of nth order equals $n!$ (in general. The number of terms containing one given element [or two given elements] is $(n-1)!$ [or $(n-2)!$ etc.]

30. You can toss a biased coin with head probability p upto 7 times. You win Rs 1000.00 if three tails appear before a head. What are your chances of winning ?

Reason has always existed, but not always in a reasonable form. (Karl Marx)

It is a mistake to confound strangeness with mystery.

(Sherlock Holmes : A Study in Scarlet)

Probabilistic Independence. Baye's Reversal Rule

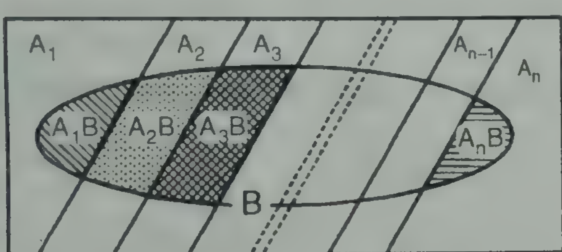
2

We consider some special rules to evaluate probabilities of compound events. We commence with the most frequently used Basic Rule :

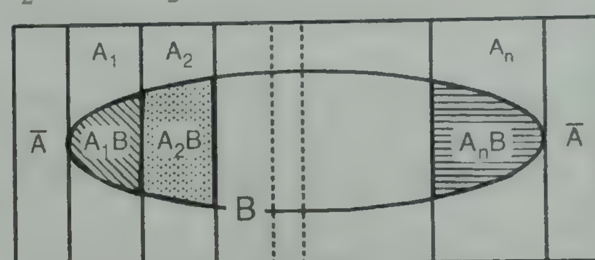
2-10. *Multi-Stage p-Rule or Theorem of Total Probability*

If $A_1, A_2, \ldots, A_n$ is a partition of the sample space Ω and B is some *proper* but arbitrary event caused by A_j, then

$$P(B) = P(A_1)\, P(B \mid A_1) + P(A_2)\, P(B \mid A_2) + \ldots + P(A_n)\, P(B \mid A_n).$$



(*i*)



(*ii*)

Proof. The events $A_1, A_2, \ldots, A_n$ form a partition of Ω, hence they are disjoint and accordingly the events $A_1B, A_2B, \ldots, A_nB$ are also disjoint [See Fig. (i)] and so by point-set algebra

$$B = A_1B \uplus A_2B \uplus \ldots \uplus A_nB.$$

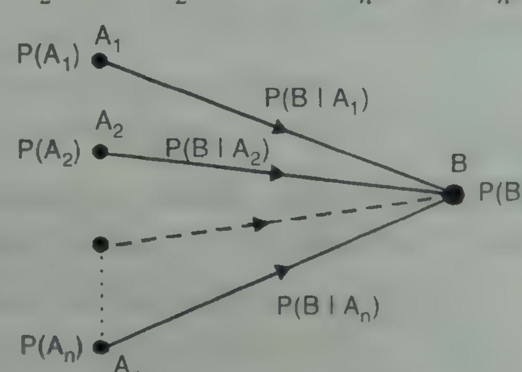
Taking probabilities of both sides and using [$P3'$] (finite additivity) we get

$$P(B) = P(A_1B) + P(A_2B) + \ldots + P(A_nB) = P(A_1)\, P(B \mid A_1) + P(A_2)\, P(B \mid A_2) + \ldots + P(A_n)\, P(B \mid A_n),$$

the latter being the consequence of the Product Rule : $P(CD) = P(C)\, P(D \mid C)$.

Special Case. Suppose $A\ (\neq \Omega)$ and B are events such that $B \subseteq A$ and $A = A_1 \uplus A_2 \uplus \ldots \uplus A_n$. Then

$$P(B) = \sum_{i=1}^{n} P(A_i)\, P(B \mid A_i),\ i = 1, 2, \ldots, n.$$



Tree diagram for the Multistage *p*-rule.

Proof. Here $\Omega = A_1 \uplus A_2 \uplus \ldots \uplus A_n \uplus A,\ B \cap \bar{A} = \varnothing$. Since, by point-set algebra. $B = A_1B \uplus A_2B \uplus \ldots \uplus A_nB \uplus \bar{A}B$, finite additivity [$P3'$] gives

$$P(B) = P(A_1B) + P(A_2B) + \ldots + P(A_nB) + P(\varnothing), \qquad \{\because \bar{A}B = \varnothing\}$$

$$= P(A_1)\, P(B \mid A_1) + P(A_2)\, P(B \mid A_2) + \ldots + P(A_n)\, P(B \mid A_n). \quad [\because P(CD) = P(C)\, P(D \mid C)]$$

Cor. $\Sigma P(B \cup A_i) = 1 + (n-1) P(B), \quad 1 \le i \le n.$

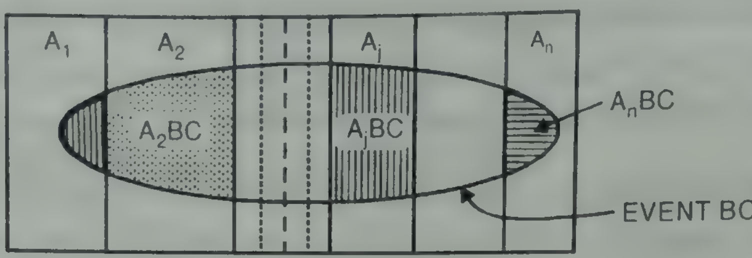
Proof. $P(A_i \cup B) = P(A_i) + P(B) - P(A_i \cap B)$, hence

$$\Sigma P(A_i \cup B) = \Sigma P(A_i) + nP(B) - \Sigma P(A_i B) = 1 + nP(B) - P(B) = 1 + (n-1) P(B).$$

2-11. *Multi-Stage Conditional p-Rule*

Let $A_1, A_2, ..., A_n, ...$ be a partition of the sample space Ω and $P(C) > 0$. Then for any event B.

$$P(B \mid C) = P(A_1 \mid C)\, P(B \mid B_1 C) + P(A_2 \mid C)\, P(B \mid A_2 C) + ... + P(A_n \mid C)\, P(B \mid A_n C).$$



Proof. The partition of BC by $\{A_j\}$ gives

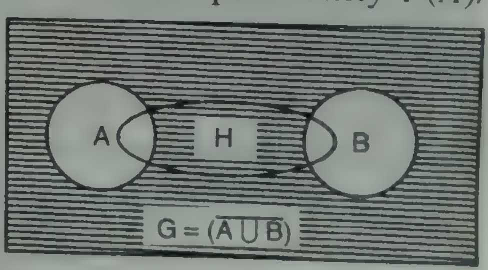
$$P(BC) = \sum_{i=1}^{n} P(A_j\, BC) = \sum_{j=1}^{n} P(C)\, P(A_j B \mid C)$$

Using $\quad P(BC) = P(C)\, P(B \mid C)$ and cancelling $P(C) \ne 0$ in the above result yields

$$P(B \mid C) = \sum_{j=1}^{n} P(A_j B \mid C) = \sum_{j=1}^{n} P(A_j \mid C) . P(B \mid A_j C).$$

2-12. *Worked-out Problems*

Example 1. Craps Rule. Let A and B be m.e. events of an experiment. If indep. trials of this experiment are performed, show that A occurs before B with probability $P(A)/[P(A) + P(B)]$.



Solution. Write $H = \{A$ occurs before B does$\}$ and

$E^{(k)} = \{$event E occurred on the kth trial$\}$

$$G = \overline{(A \uplus B)}, \quad P(G) = P(\overline{A \uplus B}) = 1 - P(A) - P(B).$$

Here $\quad S = A \uplus B \uplus G$. Now by multistage p-Rule.

$$P(H) = P(A)\, P(H \mid A^{(1)}) + P(B)\, P(H \mid B^{(1)}) + P(G)\, P(H \mid G^{(1)}) \qquad ...(1)$$

Now $\quad P(H \mid A^{(1)}) = 1,\ P(H \mid B^{(1)}) = 0,\ P(H \mid G^{(1)}) = P(H). \qquad ...(2)$

Note that, when neither A occurs nor B occurs on trial 1, the game starts afresh and Eq. 2 (iii) follows, a starting-over argument. Now substituting from (2) into (1) gives

$$P(H) = P(A) . 1 + 0 + \{1 - P(A) - P(B)] . P(H)$$

$$\therefore \quad P(H) = P(A)/[P(A) + P(B)].$$

Example 2. A man, condemned to death, is given two similar jars, 100 green balls and 100 red balls. He can distribute the balls in the two jars as he wishes. Afterwards, he is blind folded and asked to randomly choose one of the jars and then blindly draw one ball from the chosen jar. If he draws a green ball, he is pardoned and if he draws a red ball, he is executed instantly.

How this man should distribute the balls in the jars so that extraction of a green ball has maximal probability ?

Solution. Write J_1, J_2 for two jars and G denote extraction of a green ball. Suppose he puts x green balls in J_1 and $(200 - x)$ balls in J_2. Then, blind foldedly, by multistage p-Rule.

$$P(G) = P(J_1)\, P(G \mid J_1) + P(J_2)\, P(G \mid J_2)$$

$$p_x = \frac{1}{2}.1 + \frac{1}{2}.\frac{100 - x}{200 - x}.$$

Note that, derivative (dp/dx) does not help solve the problem.

Now, $\qquad p_x = (300 - 2x)/2(200 - x)$

$$p_1 = \frac{149}{199} = 0.7487,\ p_2 = \frac{148}{198} = 0.7475,\ p_3 = \frac{147}{197} = 0.7462.$$

The sequence $\langle p_n \rangle$ indicates that p_x decreases x increases. Hence, this man should put only one green ball in one jar and 199 balls in the other jar to maximize the chances of escaping death.

***Example* 3. Gambler's ruin problem : Timed play (One-unit bet).** Gambler A, who at each play of the game has chance p of winning one unit and chance $q = 1 - p$, of losing one unit. The successive plays of the game are assumed independent. Find the probability that, starting with i units, A's fortune will reach N, before reaching zero (ruinous stage).

Solution. Let $\qquad E_i = \{\text{Starting with } i \text{ units, } A \text{ eventually reaches } N\}$

$$W_1 = \{A \text{ wins the first game}\}$$

Conditioned on first game the multistage p-Rule gives

$$P(E_i) = P(W_1)\, P(E_i/W_1) + P(\overline{W}_1)\, P(E_i \mid \overline{W}_1) \qquad \text{...(i)}$$

Now $\qquad P(W_1) = p,\ P(\overline{W}_1) = q,\ P(E_i \mid W_1) = P(E_{i+1}),\ P(E_i \mid \overline{W}_1) = P(E_{i-1})$

Since winning increases assets to $(i + 1)$ units and losing decreases assets to $(i - 1)$ units and the games if continued, is as to begin with. We write

$$P(E_i) = P_i \text{ and (i) yields}$$

$$= p.P(E_{i+1}) + qP(E_{i-1}) \Rightarrow P_i = p\, P_{i+1} = qP_{i-1}$$

We write $\qquad P_i = (p + q)\, P_i$ and rewrite above as

$$p(P_{i+1} - P_i) = q(P_i - P_{i-1}) \Rightarrow P_{i+1} - P_i = \lambda\, (P_i - P_{i-1}), \qquad \text{...(1)}$$

where $\lambda = q/p$ = (lose/win) ratio. Since $P_0 = 0$ (no asset, no win), relation (1) gives

$$P_2 - P_1 = \lambda(P_1 - P_0) = \lambda P_1$$

$$P_3 - P_2 = \lambda(P_2 - P_1) = \lambda^2 P_1$$

$$P_4 - P_3 = \lambda(P_3 - P_2) = \lambda^3 P_1$$

$$P_i - P_{i-1} = \lambda\, (P_{i-1} - P_{i-2}) = \lambda^{1-i}\, P_1$$

$$\text{....} \quad \text{..........} \quad \text{.........} \quad \text{..........}$$

$$P_N - P_{N-1} = \lambda(P_{N-1} - P_{N-2}) = \lambda^{N-1}\, P_1$$

Adding the first $(i - 1)$ equations give

$$P_i - P_1 = (\lambda + \lambda^2 + \dots + \lambda^{i-1})\, P_1 \Rightarrow P_i = (1 + \lambda + \lambda^2 + \dots + \lambda^{i-1}) P_1$$

$$\therefore \quad P_i = \{(1 - \lambda^i)/(1 - \lambda)\}\, P_1, \text{ if } \lambda \neq 1, \text{ i.e. } p \neq q, \text{ i.e. } p \neq \tfrac{1}{2}\} \qquad \dots(2)$$

$$= iP_1, \text{ if } \lambda = 1, \text{ i.e. } p = q = \tfrac{1}{2}.$$

Since $\quad P_N = 1$ (Certainty), Eq. (2) using $i = N$, gives

$$P_1 = (1 - \lambda)/(1 - \lambda^N),\ p \neq \tfrac{1}{2}\ ;\ P_1 = 1/N \text{ if } p = \tfrac{1}{2}.$$

Substituting this value of P_1 in Eq. (2) yield

$$P_i = \begin{cases} (1-\lambda^i)/(1-\lambda^N), & p \neq \frac{1}{2} \qquad (\lambda = q/p) \\ i/N, & \text{if } p = \frac{1}{2} \end{cases}$$

As $N \to \infty$, $\lambda \to 0$ (if $p > \frac{1}{2}$). Then

$$P_i \to \begin{cases} (1-\lambda^i), \text{ if } & p > \frac{1}{2} \qquad (\lambda = q/p) \\ 0, \quad \text{if} & p \leq \frac{1}{2} \end{cases}$$

Consequently, if $p > \frac{1}{2}$, A's fortune increases indefinitely, but if $p < \frac{1}{2}$, A will go broke against an infinitely rich adversary.

Comments. In the Theory of Markov Chain, this problem is diversified embracing a large class of problems.

***Example* 4.** Urn A contains a white and b black balls and urn B contains C white and α black balls. A certain number $k(\leq a + b)$ of balls chosen at random is transferred from A to B. Find the probability of drawing a white ball from B after the transfer.

Solution. Among the k balls chosen from A, let X be the number of white balls so that the possible values of X are 0, 1, 2, ..., k, write

$$W = \{\text{White ball drawn from } B \text{ after the transfer of } k \text{ balls from } A\}$$

Now, by Multi-stage p-Rule

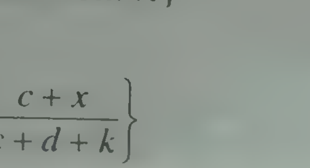
$$P(W) = \sum_{x=0}^{k} P\{X = x\}.P\{W \mid X = x) = \sum_{x=0}^{k} \left\{\binom{a}{x}\binom{b}{k-x} \Big/ \binom{a+b}{k}\right\}\left\{\frac{c+x}{c+d+k}\right\}$$

$$= \frac{1}{(c+d+k)}\left\{c \cdot \sum_{x=0}^{k} \left\{\binom{a}{x}\binom{b}{k-x} \Big/ \binom{a+b}{k}\right\}\right\} + \sum_{x=1}^{k}\left\{x \cdot \binom{a}{x}\binom{b}{k-x} \Big/ \binom{a+b}{k}\right\}$$

$$= \frac{1}{(c+d+k)}\{c \times 1 + E(X)\}, \quad X \sim HG\,(a + b, k, p)$$

$$E(X) = kp = \frac{ka}{(a+b)} = \frac{1}{c+d+k}\left(c + \frac{ka}{a+b}\right)$$

$$= \frac{[a(c+k) + bc]}{[(a+b)(c+d+k)]}.$$



Remark. Considering transfers of three balls of various combinations would require *Four* conditioning events. The above is certainly a shorter solution requiring only *two* conditioning events.

***Example* 5.** Each of the n urns contains a white balls and b black balls. One ball is transferred from the first urn into the second, then one ball from second into the third and so on. If P_k is the probability of drawing a white ball from the kth urn, show that,

$$P_{k+1} = \frac{a+1}{a+b+1} P_k + \frac{a}{a+b+1}(1 - P_k).$$

Deduce that for the last urn, $P_n = a/(a + b)$.

Solution. Let us denote kth urn by U_k. A white ball from U_k can be extracted in the following two disjoint ways :

(i) $A = \{$Ball transferred from U_{k-1} to U_k is white$\}$, $P(A) = P_{k-1}$.

(ii) $B = \{$Ball transferred from U_{k-1} to U_k is black$\}$, $P(B) = 1 - P_k - 1$.

$W = \{$a white ball is extracted from U_k after the transfer$\}$,

So $\quad P(W) = P(A)\, P(W \mid A) + P(B)\, P(W \mid B)$ [Multi-stage p-Rule]

$$P_k = P_{k-1} \cdot \frac{a+1}{a+b+1} + (1 - P_{k-1}) \cdot \frac{a}{a+b+1} \quad \text{[with obvious evaluations]}$$

i.e. $\quad P_k = cP_{K-1} + ac, \quad c = 1/(1 + a + b)$ ...(1)

$$\begin{aligned} P_k &= ac + cP_{k-1} = ac + c\{ac + cP_{k-2}\} = ac + ac^2 + c^2 P_{k-2} \\ &= (ac + ac^2) + c^2[ac + c\,P_{k-3}] = ac + ac^2 + ac^3 + c^3\, P_{k-3} = \ldots \\ &= ac + ac^2 + ac^3 + \ldots + ac^{k-1} + c^{k-1} P_1 \qquad [P_1 = a/(a = b)] \\ &= ac(1 + c + c^2 + \ldots + c^{k-2}) + c^{k-1} \,.\, [a/(a = b)] \\ &= ac \,.\left(\frac{1 - c^{k-1}}{1 - c}\right) + \frac{ac^{k-1}}{a+b}\left[c = \frac{1}{1+a+b}, 1 - c = \frac{a+b}{a+b+1} = (a+b)\,c\right] \\ &= \frac{a}{a+b}\{1 - e^{k-1} + e^{k-1}\} = \frac{a}{a+b}. \end{aligned}$$

Remark. The solution of difference equation $P_n = xP_{n-1} + y$, is given by

$$P_n = (P_1 - \lambda)\, x^{n-1} + \lambda, \; \lambda = y/1(1 - x).$$

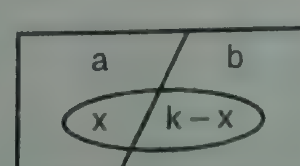
This provides result immediately from (1) : an excellent alternative solution.

***Example* 6.** From an urn containing a white balls and b black balls, a certain number of k balls is drawn, and they are laid aside, their colour unnoted. Then one more ball is drawn. Find the chance that it is a white or a black ball.

Solution. Let X be the number of white balls among the k removed balls and Y be the number of white balls that remained in the urn so that $X + Y = a$. Let $W = \{$ball drawn, after the removal of k balls, is white$\}$. Then, by Multi-stage p-Rule.

$$P(W) = \Sigma P\{W \mid Y = 0\}\; P(Y = y)$$

$$= \Sigma \frac{y}{(a+b-k)} P(Y = y)$$



$$= \frac{1}{(a+b-k)}\left\{\sum_y yP(Y=y)\right\} = \frac{1}{(a+b-k)} E(Y)$$

From $X + Y = a$, we get $E(X + Y) = a \Rightarrow E(Y) = a - E(X)$.

As $X \sim HG\ (a + b, a; k)$ so that $E(X) = kp = k\ .\ \{a/(a + b)\}$

$$\therefore \quad P(W) = \left\{a - \frac{ka}{a+b}\right\} \Big/ (a+b-k) = \frac{a(a+b-k)}{(a+b)(a+b-k)} = \frac{a}{a+b}.$$

***Example* 7.** The game called **craps** is played with two dice. The caster wins at once if the sum is 7 or 11 (called *naturals*). He loses at once if the sum is 2, 3 or 12 (called *craps*). But if he produces sums 4, 5, 6, 7, 8, 9 or 10 (called *points*) he has the right to cast the die repeatedly until he throws his *point* again and wins or loses by throwing a seven. Find the probability of caster's winning.

Solution. Let W_1 denote that the player wins on the first throw and G denote that the player wins on game points condition. If W denotes that the player wins, then

$$P(W) = P(W_1 \uplus G) = P(W_1) + P(G) \qquad \text{...(1)}$$

For a pair of dice, sum k, we have

$$P(S_k) = \{6 - |7 - k|\}/36 \qquad \text{...(i)}$$

Also $\quad P\{A \text{ wins before } B\} = P(A)/[P(A) + P(B)] \qquad$ ...(ii)

Now $\quad P(W_1) = P\{S_7 \uplus S_{11}\} = P(S_7) + P(S_{11}) = \dfrac{6}{36} + \dfrac{2}{36} = \dfrac{2}{9} = 0.2222 \qquad$ ...(2)

Player's game plants are $T = \{4, 5, 6, 8, 9, 10\}$

$$P(S_4) = P(S_{10}) = 3/36, \quad P(S_5) = P(S_9) = 4/36, \quad P(S_6) = P(S_8) = 5/36.$$

Further by (ii),

$$P(G \mid S_k) = P\{\text{Game won given } S_k \text{ occurs } S_7,\ k \in T\}$$

$$= \frac{P(S_k)}{P(S_k) + P(S_7)} = \frac{6 - |7 - k|}{2 - |7 - k|} \qquad \text{[Using (i)]}$$

$$\therefore \quad P(G) = \sum_{k \in J} P(S_k)\, P(G \mid S_k) \qquad \text{[by Multi-stage } p\text{-Rule]}$$

$$= \sum_{k \in J}\left\{\frac{6 - |7 - k|}{36} \cdot \frac{6 - |7 - k|}{12 - |7 - k|}\right\} = \frac{1}{36}\sum_T \frac{\{6 - |7 - k|\}^2}{12 - |7 - k|}$$

$$= \frac{2}{36}\left(1 + \frac{8}{5} + \frac{25}{11}\right) = \frac{134}{495} = 0.270707. \qquad \text{...(3)}$$

Substituting from (2) and (3) into (1) yields

$$P(W) = 0.222222 + 0.270707 = 0.4293 = 0.493.$$

***Example* 8.** Persons A and B often forget their umbrellas in some way or another. A always carries his umbrella when he goes out while B does it 50% of the time. Both A and B forget their umbrellas at any shop with probability p. After visiting their shops, they return home. Find

(a) $P\{$They both have umbrellas$\}$, (b) $P\{$They have only one umbrella$\}$,
(c) $P\{B$ has lost his umbrella given that there is only one umbrella after their return$\}$.

Solution. We designate the events as under :

$A = \{A$'s umbrella intact$\}$, $A_j = \{A$ does not forget his umbrella in shop $j\}$, $H = \{B$ has left his umbrella at home$\}$

$E_1 = \{$They have only one umbrella$\}$. Terminology for B is similar.

$$P(AB) = P(A_1A_2A_3) \cap P(\overline{H}B_1B_2B_3) \uplus (A_1A_2A_3)H\}, \qquad \text{[At home after shopping]}$$

$$= P(A_1A_2A_3)\ P(\overline{H})P(B_1B_2B_3) + P(A_1A_2A_3)\ P(A) = q^3\left(\tfrac{1}{2}\right)q^3 + q^3\left(\tfrac{1}{2}\right) = \tfrac{1}{2}q^3(1+q^3)$$

$$P(\overline{A}) = P(A \text{ forgets his umbrella}) = P(\overline{A}_1 \uplus A_1\overline{A}_2 \uplus A_1A_2\overline{A}_3) = p + qp + q^2p = p(1+q+q^2)$$

$$P(\overline{B}) = P(B \text{ forgets his umbrella}\} = P\{\overline{H}\overline{B}_1 \uplus \overline{H}B_1\overline{B}_2 \uplus \overline{H}B_1B_2\overline{B}_3\}$$

$$= \tfrac{1}{2}p + \tfrac{1}{2}qp + \tfrac{1}{2}q^2p = \tfrac{1}{2}p(1+q+q^2)$$

$$P(A) = 1 - P(\overline{A}) = 1 - p(1+q+q^2) = q^3$$

$$P(B) = 1 - P(\overline{B}) = 1 - \tfrac{1}{2}p(1+q+q^2) = \tfrac{1}{2}(1+q^3)$$

$\{$Only one umbrella$\} = A\overline{B} \uplus \overline{A}B$

$$\therefore \quad P(A\overline{B} \uplus \overline{A}B) = P(A)\,P(\overline{B}) + P(\overline{A})\,P(B)$$

$$= \tfrac{1}{2}p(1+q+q^2).q^3 + \tfrac{1}{2}(1+q^3)\,p(1+q+q^2)$$

$$= \tfrac{1}{2}p(1+q+q^2) + (1+2q^3) \ : \text{[Probability of only one umbrella]}$$

(c) $P\{A\overline{B} \mid (A\overline{B} \uplus \overline{A}B)\} = P\{A\overline{B}\}/P\,(A\overline{B} \uplus \overline{A}B)$

$$= \tfrac{1}{2}p(1+q+q^2).q^3 \big/ \tfrac{1}{2}p(1+q+q^2)(1+2q^3)$$

$$= q^3/(1+2q^3)$$

Problems with Solutions Provided at the End of the Text

1*. **Between-ness property.** For any event B, the value of $P(A)$ lies between those of $P(A \mid B)$ and $P(A \mid \overline{B})$.

2*. Two cards are drawn at random sucessively without replacement from a full pack of 52 playing cards. Find the probability that the second card may **cover** the first card. [That is, the 2nd card must be *a superior* card of the same suit].

3*. Find the probability that the ace of spades (A_S) is next to the king of spades (K_S) in a well-shuffled pack of 52 cards.

4*. In a group of equal no. of men and women, 10% men and 45% women are unemployed. What is the prob. that a person selected at random is employed ?

5*. An urn contains 6 tickets numbered 1, 2, 3, 4 5, 6. Another urn contains 5 tickets numbered 2, 4, 6, 8, 10. One urn is chosen at random and a ticket is drawn from it. Find the probability that the ticket drawn bears the number :

(a) 3 or 6, (b) 4 or 6, (c) 1 or 10, (d) 5, (e) 10.

6*. An urn contains 3 red and 6 black balls, and a ball is chosen at random and set aside *without noting* its colour. If a second ball is now withdrawn, what is the probability that it is red ? How is the probability changed, if we know the colour of the first ball ?

7*. A box contains 4 balls numbered 0 through 3. One ball is chosen at random and kept aside. All balls with non-zero numbers less than that of the kept-aside ball are also removed from the box. Now a 2nd ball is chosen at random from those remianing in the box. Find the probability that the 2nd ball chosen is numbered 3.

8*. An urn contains N balls, among them just W are white. A random sample of size n is drawn without replacement and from this sample another random sample of size m is drawn without replacement. Find the chance that the second sample contains just k white balls.

9*. If the probability that a family has exactly n children is $(1/2)^n$, $n = 1, 2, 3, \ldots$ and if all 2^n permutations of the sexes of the n children are eqully likely, find the probability that a family will have no boys.

10*. (*Laplace Rule of Succession*) of n biased coins, with probability (k/n) of falling heads for the kth coin, a coin is chosen at random and tossed $(a + b)$ times. What is the probability that the last b tosses of the coin result in heads, when the first a tosses have already resulted in heads ?

11*. Let the probability P_n that a family has n children be αp^n, and $P_0 = 1 - \alpha p\,(1 + p + p^2 + \ldots)$ $n \geq 1$. Suppose that all sex distributions of n children have the same distribution.

(a) Show that for $k \geq 1$, the probability that a family contains exactly k boys is $2\alpha p^k/(2 - p)^{k+1}$.

(b) Given that a family includes at least one boy, show that the probability that there are two or more boys is $p/(2 - p)$.

Exercise 2(a)

1. Comment on the statement (Simpson's paradox) : It is possible to have $P(A \mid B) < P(A \mid B)$, even though $P(A \mid BC) \geq P(A \mid \bar{B}C)$; $P(A \mid B\bar{C}) \geq P(A \mid \bar{B}\bar{C})$.
2. In a group of employed persons 28% are women. 65% of the men and 40% of the women pay income tax. Show that the probability, that a randomly chosen person is a non-income-tax payer is 0.42.
3. A bag contains 3 white and 2 black balls, and another contains 5 white and 3 black balls. If a bag is chosen at random and a ball is drawn from it, show that the probability that it is white is 49/80.
4. Two-thirds of the students in a class are boys and the rest are girls. It is known that the probability of a girl getting a first class is 0.25 and that of a boy getting first class is 0.28. Prove that the probability that *a* student chosen at random will get first class marks in the subject is 0.27.
5. Next year, there will be three candidates Mr. A, Mr. B, Mr. C for the position of a principal, whose respective chances of getting appointment are in proportion 4 : 2 : 3. The probabilities that these

persons, if selected, will introduce co-education in the college, are respectively 0.3, 0.5, 0.8. Show that the probability that there will be co-education in the college next year is 23/45.

6. A fair coin is flipped. If a head appears, then a symmetrical tetrahedral die (with faces marked 1, 2, 3, 4) is rolled. If a tail appears, then an ordinary cubic die is rolled. Show that probability that the number of points on the die will be more than two is 7/12.

7. There are five guns that, when properly aimed and fired, have respective probabilities of hitting the target as follows : 0.5, 0.6, 0.7, 0.8 and 0.9. One of the guns is chosen at random, aimed and fired. Prove that the probability that the target is hit is 7/10.

8. A certain drug manufactured by a company is tested chemically for its toxic nature. Event "drug is toxic" is denoted by T and the event "the chemical test reveals that the drug is toxic" is denoted by C. If $P(T) = \theta$, $P(C \mid T)$ $(\bar{C} \mid \bar{T}) = 1 - \theta$, show that the probability that the drug is not toxic, given that the chemical test reveals that it is toxic, is free from θ.

9. H is one of the six horses entered for a race and is to be ridden by one of the two jockeys J_1 or J_2. It is 2 : 1 that J_1 rides H, in which case all the horses are equally likely to win; if J_2 rides H, his chance is trebled. Show that the odds against his winning are 13 : 5.

10. One and only one of A, B, C, D has committed a murder. On the basis of initial information any one of them had the same chance of committing it. None of them would plead guilty if in fact he was not guilty. Their respective chances of pleading guilty, having committed the crime, are judged to be 0.5, 0.4, 0.3, 0.2. All of them have pleaded not guilty. What is the probability that D did it. ? **[Ans.** $p = 4/13$]

11. An urn contains a white balls and b black balls, another contains c white balls and d black balls. One ball is transferred from the first urn into the second, and then a ball is drawn from the latter. Show that the chance that it will be a white ball is $[a(c + 1) + bc]/(a + b)(c + d + 1)$.

12. An urn A contains 3 white and 4 black balls and an urn B contains 5 white and 3 black balls. One ball is chosen randomly from A and transferred to B. Two balls are chosen without replacement from B. Show that the chance that they are of the same colour is 59/126.

13. Urn A contains 6 white and 4 black balls and urn B contains 2 white and 2 black balls. From urn A two balls are transferred to B and a sample of size 2 is then drawn without replacement from urn B. What is the probability that the sample contains (i) Exactly 1 white ball ? (ii) Both balls are black ? (iii) Balls are of the same colour ? **[Ans.** $128\lambda, 41\lambda, 97\lambda, \lambda = 1/225$]

14. Urn A contains 10 white and 3 black balls, urn B contains 3 white and 5 black balls. Two balls are transferred from A to B and then one ball is drawn from B. What is the probability that it is (i) white ball, (ii) black ball ? **[Ans.** 59/130, 71/130]

15. A box contains n_1 tags numbered n_1 and n_2 tags numbered 2. A tag is selected at random. If it is No. 1 tag, one goes to urn A which contains r_1 red balls and b_1 black balls and selects a ball at random; if it is No. 2 tag, one goes to urn B which contains r_2 red balls and b_2 black balls, and selects a ball at random. Find the probability that one obtains a red ball.

16. The probability that a certain beginner at gold gets a good shot if he uses the correct club is 1/3, and the probability of a good shot with an incorrect club is 1/4. In his bag are five different clubs, only one of which is correct for the shot in question. If he chooses a club at random and takes a stroke, show that the probability that he gets a good shot is 4/15.

17. In a clubroom, there are 5 ordinary bridge decks and 3 pinochle decks, all having similar construction and designs. One of these 8 decks is chosen at random, and a card is randomly drawn from it. If the card is jack of hearts, find the chance that it came from (i) a Pinochle deck, (ii) a Bridge deck. **[Ans.** $p_1 = 13/23$, $p_2 = 10/23$]

18. Urn A contains a white and b black balls, urn B contains b white and a black balls. Starting with the first drawing from A, a series of drawings is made according to the following rules :

(i) Each time only one ball is drawn and immediately returned to the same urn it came from.

(ii) If the ball drawn is white, the next drawing is made from A.

(iii) If the ball drawn is black, the next drawing is made from B.

What is the chance that kth ball drawn will be white ? $\left\{p_k = \frac{1}{2} + \frac{1}{2}[(9a - b)/(a+b)]^k\right\}$.

19. (a) From a box containing 4 black and 6 red balls, 5 balls are transferred into an empty urn. From the urn, 3 balls are drawn and they happen to be black. Show that the probability that the fourth ball taken from the urn will also be black is 1/7.

(b) There are 3 boxes containing red and white chips with r_j red chips and w_j white chips in Box j, $1 \leq j \leq 3$. Two fair cons are tossed and one chip is drawn from the box corresponding to the number of heads showing plus 1. Prove that

$$P\{\text{chips is red}\} = (r_1/4c_1) + (r_2/2c_2) + cr_3/4c_3), \quad [c_j = r_j + w_j].$$

20. Two electric supply units A and B operate in parallel to meet the power requirements of a small city. Each unit has a capacity so that it can supply the city's full power requirements 75% of the time, in case the other unit fails. The probability of failure of each unit is 0.15 that of both units failing 0.03. If there is failure in the power generation, find the probability that the city has its suppy of full power.

2.20. *Independence of Events*

Events A and B are said to be independent, [written : $I(A, B)$ or Ind (A, B)], iff

$$P(AB) = P(A)\,P(B).$$

Independence of three events A, B, C is defined by four relations.

$P(AB) = P(A)\,P(B)$; $P(AC) = P(A)\,P(C)$, $P(BC) = P(B)\,P(C)$; $P(ABC) = P(A)\,P(B)\,P(C)$.

A family of events is completely independent iff its every fintie subcollection of events is independent.

In general, events $A_1, A_2, \ldots, A_n$ are said to be *pairwise independent* iff

$$P(A_iA_j) = P(A_i)\,P(A_j), \text{ for all } i \neq j.$$

Events $A_1, A_2, \ldots, A_n$ are said to be *mutually independent* iff

$$P(A_iA_j) = P(A_j)\,P(A_j), \ \forall\, i, j, i \neq j.$$
$$P(A_iA_jA_k) = P(A_i)\,P(A_j)\,P(A_k), \ \forall\, i, j, k, i \neq j \neq k.$$
$$\ldots \ldots \ldots \ldots \ldots \ldots \ldots \qquad \ldots(1)$$
$$P(A_1A_2, \ldots, A_n) = P(A_1)\,P(A_2), \ldots, P(A_n).$$

Precisely, events A_i are said to be independent iff for each finite set of distinct indices $i_1, i_2, \ldots, i_k \in A$ (index set), we have

$$P(A_{i_1}, A_{i_2}, \ldots A_{i_k}) = P(A_{i_1})\,P(A_{i_2}) \ldots P(A_{i_k}).$$

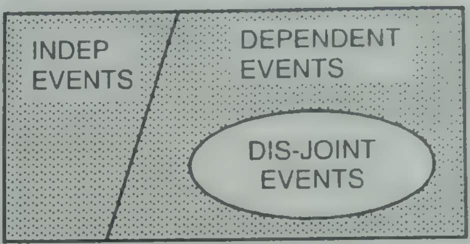
Abbreviations. Sometimes the words, "independent" and "dependent" shall be written "indep" and "dep" respectively.

2-21. Number of Conditions for Mutually Independent Events

If n events are completely independent, then every combination of n events, taken any number of times, should be independent. It follows that, the number of relations in the $(n-1)$ equations listed in (1) have $\binom{n}{2}, \binom{n}{3}, \binom{n}{4}, \ldots, \binom{n}{n}$ terms respectively. Thus, the total number of conditions among the relations (1) are

$$\binom{n}{2}+\binom{n}{3}+\ldots+\binom{n}{n} = (1+1)^n - \binom{n}{0} - \binom{n}{1} = 2^n - 1 - n.$$

2-22. Independence versus Mutual Exclusiveness



(i) A and B are (proper) mutually exclusive events and $P(A) \cdot P(B) > 0$, then they cannot be independent.

(ii) If A and B are (proper) independent events and $P(A) \,.\, P(B) > 0$, then they cannot be mutually exclusive.

Proof. We are given that $P(A) \neq 0$, $P(B) \neq 0$, i.e. $P(A) \,.\, P(B) > 0$.

(i) If events A and B are disjoint, $AB = \varnothing$ and so $P(AB) = 0 \neq P(A) \,.\, P(B) > 0$.

It follows that, A and B cannot be independent events.

(ii) If A and B are independent events, then $P(A \cap B) = P(A)\, P(B) > 0$.

Thus $P(A \cap B) \neq 0$. Hence A and B cannot be disjoint.

Comments. Events A and B have to be proper $[P(A) \neq 0, P(B) \neq 0]$, otherwise $A = \Omega$, $B = \varnothing$, $\varnothing \cap \Omega = \varnothing$ so $P(\varnothing \cap \Omega) = P(\varnothing)\, P(\Omega) = 0$, is true, so that disjoint events are independent.

Further Ind $(A, B) \not\Rightarrow P(A \mid B) = P(A)$, because conditional probability is not defined for $B = \varnothing$.

Remarks. Events can be dependent without being disjoint. In fact, the disjoint sets of events can be depicted as a subset of the dependent sets of events of the sample space. [See figure]. Independence and mutually exclusiveness may be persent in the same problem :

Illustration. Define the events for a single roll of a die : $A = \{1, 3, 5\}$, $B = \{2, 4, 6\}$, $C = \{5, 6\}$

Then, $\quad A \cap B = \varnothing$, $B \cap C = \{6\} \neq \varnothing$. Now,

(i) $P(AB) = 0 \neq P(A)\, P(B) = 1/4$. Thus A and B are disjoint but not independent.

(ii) $P(B \cap C) = \frac{1}{6} = \frac{1}{2} = \frac{1}{3} = P(B) \,.\, P(C)$. Thus B and C are independent but not disjoint.

Example. A single card is drawn from an ordinary deck. Give examples of events A and B associated with this experiment that are

(a) mutually exclusive (m.e.) but not independent (indep.) (b) indep. but not m.e.

(c) indep. and m.e. (d) neither indep. nor m.e.

Solution. We define the following events :

$S = \{\text{Spade is obtained}\}$, $H = \{\text{Heart is obtained}\}$, $A = \{\text{Ace is obtained}\}$, $D = \{S \cup H\}$

(a) $P(S \cap H) = 0 \neq P(S)\, P(H)$, (b) $P(S \cap A) = 1/52 = P(S)\, P(A) = [(1/13)\,(1/4)]$

(c) Class of m.e. and indep. events is disjoint.

(d) $P(S \cap D) = P(S) = 1/4$, $P(S) \,.\, P(D) = (1/4)\,(1/2) = 1/8$.

2-23. Some Theorems Regarding Independence

***Theorem* 1.** If A and B are independent events, then

(i) A and $\overline{B}$ are independent. (ii) $\overline{A}$ and B are independent, (iii) $\overline{A}$ and $\overline{B}$ are independent.

Proof. Since A and B are independent, we have $P(A \cap B) = P(A)\ P(B)$ ...(1)

Also $P(C \cap \overline{D}) = P(C) - P(CD)$. [Subtractive law] ...(2)

(i) $P(A\overline{B}) = P(A) - P(AB) = P(A) - P(A)\ P(B)$ [by (1) and (2)]

$= P(A)\ [1 - P(B)] = P(A)\ .\ P(B)$ [by Complement Rule]

(ii) $P(\overline{A}B) = P(B) - P(AB) = P(B) - P(A)\ P(B).$ [by (1) and (2)]

$= P(B)[1 - P(A)] = P(B).P(\overline{A}) = P(\overline{A})\,P(B)$ [by Complement Rule]

(iii) $P(\overline{A} \cap \overline{B}) = 1 - P(A \cup B)$ [by Complement Rule]

$= 1 - [P(A) + P(B) - P(AB)]$ [Union of 2 events]

$= 1 - P(A) - P(B) + P(A)\ P(B)$ [by (1)]

$= [1 - P(A)][1 - P(B)] = P(\overline{A})\,P(\overline{B})$. [by Negation Rule]

Aliter. $P(\overline{A}\,\overline{B}) = P(\overline{A}) - P(\overline{A}\,B) = P(\overline{A}) - P(\overline{A})\,P(B)$ [by (2) and (iii)]

$= P(\overline{A})[1 - P(B)] = P(\overline{A}).P(\overline{B}),$ [by Negation Rule]

This means A and B are independent events.

Comment. Perhaps it is now obvious that $I(A\ ;\ B) \Leftrightarrow I(A;\overline{B})$ etc. Thus, the converse of theorem 1 is true, since $(A^c)^c = A$.

***Theorem* 2.** The following conditions are equivalent.

(a) A, B, C are independent, (b) $\overline{A}, B, C$ are independent.

(c) $\overline{A}, \overline{B}, C$ are independent (d) $\overline{A}, \overline{B}, \overline{C}$ are independent.

Proof. For two events A, B, Independent (A, B), Ind. $(\overline{A}, B)$, Ind. $(\overline{A}, \overline{B})$ are equivalent. This result shows that A, B, C are pairwise independent if $\overline{A}, B, C$ are pairwise independent. This gives,

$$P(\overline{A}\,BC) = P(\overline{A})\,P(B)\,P(C) \quad ...(1)$$

$$\text{L.H.S.} = P(BC - ABC) = P(BC) - P(ABC) = P(B)\ P(C) - P(ABC)$$

$$\text{R.H.S.} = [1 - P(A)]\ P(A)\ P(B) = P(A)\ P(B) - P(A)\ P(B)\ P(C)$$

So Eq. (2) yields $P(ABC) = P(A)\ P(B)\ P(C)$.

Just show equivalence $(a) \Leftrightarrow (b)$. Then $(a) \Leftrightarrow (c)$ and $(a) \Leftrightarrow (d)$ follows by repeated use of $(a) \Leftrightarrow (b)$.

***Theorem* 3.** If the events A, B, C are **completely** independent, then the pairs (A, BC), (B, CA), (C, AB) also consist of independent events.

Proof. Since A, B, C are completely independent, the following relations are valid $P(AB) = P(A)\ P(B)$, $P(BC) = P(B)\ P(C)$, $P(CA) = P(C)\ .\ P(A)$, $P(ABC) = P(A)\ P(B)\ P(C)$. ...(1)

For the pair (A, BC), we note that

$$P(A(BC)) = P(ABC) = P(A)\ P(BC).$$ [by [1(b)] and [1(d)]

Thus A and BC are independent. Others also follows trivially.

***Theorem* 4.** The event A is known to be independent of the events B, $B \cup C$ and $B \cap C$. Then A is also independent of C.

Proof. We are given that, by independence,

$$P(AB) = P(A)\,P(B),\quad P[A(B \cup C)] = P(A)\,P(B \cup C),\ P[A(BC)] = P(A)\,P(BC) \qquad \ldots(1)$$

Now $P(AB \cup AC) = P(AB) + P(AC) - P[A(BC)] = P(A)\,P(B) + P(AC) - P(A)\,P(BC)$[by (1)] ...(i)

$$P[A(B \cup C)] = P(A)\,P(B \cup C) = P(A)\,[P(B) + P(C) - P(BC)], \qquad \text{[by 1(b)]} \quad \ldots\text{(ii)}$$

Equating (i) and (ii) [because $A(B \cup C) = AB \cup AC$] we get

$$P(AC) = P(A)\,P(C) \Rightarrow A \text{ is independent of } C.$$

***Theorem* 5.** Let $A_1, A_2, \ldots, A_m$ be disjoint events and let $B_1, B_2, \ldots, B_n$ be also disjoint events. If events A_i and B_j are independent for every choice of i and j, then A and B are independent, where

$$A = \uplus_i A_i,\ B = \uplus_j B_j,\ 1 \le i \le m,\ 1 \le i \le m$$

Proof. Here $AB = \left(\bigcup_{i=1}^{n} A_i\right) \cap \left(\bigcup_{j=1}^{n} B_j\right) = \bigcup_{i=1}^{n}\bigcup_{j=1}^{n} A_i B_j$ and the events $C_{ij} = A_i B_j$ are disjoint $[C_{ij} \cap C_k l = \varnothing]$. Hence

$$P(AB) = \sum_{i=1}^{m}\sum_{j=1}^{m} P(A_i B_j) = \sum_{i=1}^{m}\sum_{j=1}^{m} P(A_i)\,P(A_j) = \left(\sum_{i=1}^{m} P(A_i)\right)\left(\sum_{j=1}^{m} P(B_j)\right) = P(A)\,P(B).$$

***Theorem* 6. Probability of at least one of n independent events :**

Let $A_1, A_2, \ldots, A_n$ be n mutually independent events, with $P(A_j) = p_j$ and $P(\bar{A}_j) = q_j$, $[q_j = 1 - p_j]$. Now by Negation Rule

$$P(L_1) = P\{A_1 \cup A_2 \cup \ldots \cup A_n\} = 1 - P\{\overline{A_1 \cup A_2 \cup \ldots \cup A_n}\} = 1 - P[\bar{A}_i \cap \bar{A}_2 \cap \ldots \cap \bar{A}_n].$$

$$= 1 - P(\bar{A}_1 . P(\bar{A}_2) \ldots P(\bar{A}_n) = 1 - q_1, q_2, \ldots q_n. \qquad [\because\ \bar{A}_i \text{ are independent}]$$

In particular if A and B are two independent events, and if $L_1 = A \cup B$, then $P\{L_1\} = 1 - P(\bar{A})\,P(\bar{B})$.

Comment. For independent events, this rule is very convenient.

2-24. *Conditional Independence of Events*

The events A and B are said to be conditionally indepedent, if given C

$$P(AB \mid C) = P(A \mid C) \,.\, P(B \mid C), \qquad [\text{or } P(A \mid BC) = P(A \mid C)].$$

The notion of conditional independence can be extended to more than two events.

Illustration. Company C_1 produces on the average one defective tube per 100 tubes and company C_2 produces only one defective tube per 200 tubes. A box received from C_1 or C_2 is opened, a tube checked and it is non-defective. Find the probability that the 2nd tube checked shall also be non-defective.

Solution. Let T_1 = {First tube is non-defective}, T_2 = {2nd tube is non-defective}. Let $C_i\,(i = 1, 2)$ denote that the box received is from company C_i. We express our aprori ignorance by setting $P(C_1) = 1/2 = P(C_2)$. Now if T denotes a non-defective tube, then given data is

$$P(\bar{T} \mid C_1) = 0.01,\ P(\bar{T} \mid C_2) = 0.005.$$

So by Multi-stage Rule

$$P(T_1 T_2) = P(T_1 T_2 \mid C_1) P(C_1) + P(T_1 T_2 \mid C_2) P(C_2) = \tfrac{1}{2}(99/100)^2 + (199/200)^2\}$$

$$P(T_1) = P(T_1 \mid C_1) P(C_1) + P(T_1 \mid C_2) P(C_2) = \tfrac{1}{2}\{(99/100) + (199/200)\}$$

$$\therefore \quad P(T_2 \mid T_1) = P(T_1 T_2 / P(T_1) = [(198)^2 + (199)^2/(198 + 199) \times 200.$$

Comments. Here, states of successive tubes are cond. independence $P(T_1 \bar{T}_2 \mid C_1 = P(T_1 \mid C_1) P(\bar{T}_2 \mid C_1)$.

These states are not independent. $P(T_2 \mid T_1) \neq P(T_2) = P(T_1)$.

The topic of conditional independence is of growing importance.

2-25. *Worked-Out Problems*

***Example* 1.** Give examples of three events which are : (i) Independent in pairs but not mutually independent, (ii) Independent three-wise but not independent.

Solution. (i) Define the events relating to throw of two dice D_1 and D_2 :

A = {even numbers on D_1}, $\quad B$ = {odd numbers on D_2}

C = {both even numbers or both odd numbers}.

$$\therefore \quad P(A) = P(B) = P(C) = \tfrac{1}{2}, P(AB) = P(BC) = P(CA) = \tfrac{1}{4}.$$

Obviously, events A, B, C are pairwise independent events. However, since $A \cap B \cap C = \varnothing$. $P(ABC) = 0$ so that $P(ABC) \neq P(AB)\, P(C)$. Thus AB is not independent of C.

(ii) Let $\Omega = \{1, 2, 3, ..., 16\}$; $P(e_i) = 1/16$; $1 \le i \le 16$, $e_i \in \Omega$ and consider the events $A = \{1, 2, 3, 4, 5, 6, 10, 15\}$, $B = \{1, 2, 3, 4\}$, and $C = \{1, 7, 8, 9, 11, 12, 13, 16\}$.

The various combinations are : $A \cap B = B$, $A \cap C = \{1\}$, $B \cap C = \{1\}$, $A \cap B \cap C = \{1\}$.

$$P(AB) = \tfrac{1}{4} \neq P(A)\, P(B); \qquad P(BC) = \tfrac{1}{16} \neq P(B)\, P(C) = \tfrac{1}{4} \cdot \tfrac{1}{2},$$

$$P(CA) = \tfrac{1}{16} \neq P(A)\, P(C); \qquad P(ABC) = \tfrac{1}{16} = P(A) . P(B) . P(C).$$

Thus, no two of A, B, C are independent although $P(ABC) = P(A)\, P(B)\, P(C)$.

***Example* 2.** (i) If A and B are independent events and $A \subseteq B$, prove that either $P(A) = 0$ or $P(B) = 1$.

(ii) If an event A is independent of itself, (i.e. A and A are independence) show that $P(A) = 0$ or $P(A) = 1$.

(iii) If $P(A) = 0$ or 1, show that A and B are independence for any event B, in particular A and A are independent.

Solution. Here : $P(A \cap B) = P(A)\, P(B) \qquad [I(A\,;\,B)]$ ...(1)

(i) Since $A \subseteq A$, we get $A \cap B = A$, and (1) provides $P(A)\, [1 - P(B)] = 0 \Rightarrow P(A) = 0$ or $P(B) = 1$.

(ii) Since $A \cap A = A$, Eq. (1) reads $P(A) = P(A)\, .\, P(A)$

so that $P(A)\, [1 - P(A)] = 0 \Rightarrow P(A) = 0$ or $P(A) = 1$.

(iii) Let $P(A) = 1$. Since $A \cup B \supseteq A$, we do have $P(A \cup B) \geq P(A) = 1$.

Since probability measure cannot exceed unity, this realy reads :

$$1 = P(A \cup B) = P(A) + P(B) - P(AB) \Rightarrow P(B) = P(AB).$$

So $\quad P(AB) = 1 . P(B) = P(A) . P(B) \Rightarrow I(A \,;\, B)$ true.

Let $P(A) = 0$. Since $AB \subseteq A$, we do have $P(AB) \leq P(A) = 0$.

Since probability measure is non-negative, this really gives

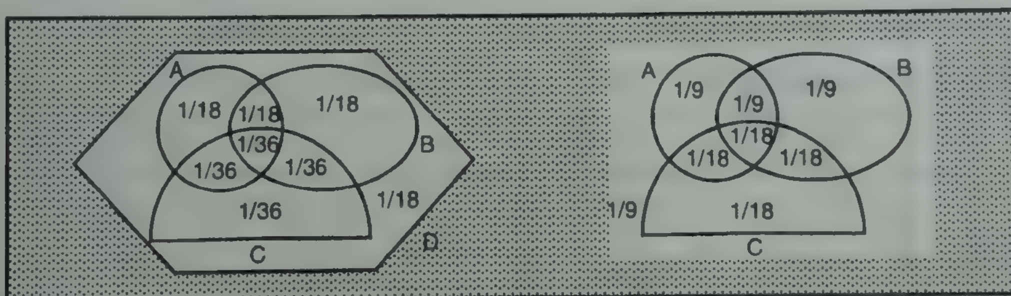
$$P(AB) = 0 \Rightarrow P(A)\,P(B) = 0. \qquad P(A)\,P(B) = P(AB).$$

This shows that A and B are independent. In particular, replacing B by A, we see that A and A are independent.

Remarks. In general, A is not independent of A. Recall that $I(A \,;\, B) \Rightarrow P(A \mid B) = P(A \mid \bar{B})$.

Hence if $B = A$, we get $P(A \mid A) = 1 \neq 0 = P(A \mid \bar{A})$. Thus, "Statistical independence" is not a "reflexive relation" on the event space.

***Example* 3.** Four events A, B, C, D, depicted in the following figure are with assigned probabilities. Show that these events are independent. [Regions representing events A, B, C consist of two subregions each separated apart. A may be 'rains' in two parts of a city with dry-spell apart].



Solution. It is convenient to label the events by subscripts so that left set is A_1, B_1, C_1 and right set is A_2, B_2, C_2. Now

$$P(A) = P(A_1) + P(A_2) = \tfrac{1}{6} + \tfrac{1}{3} = \tfrac{1}{2}$$

$$P(B) = P(B_1) + P(B_2) = \tfrac{1}{6} + \tfrac{1}{3} = \tfrac{1}{2}$$

$$P(C) = P(C_1) + P(C_2) = \tfrac{1}{9} + \tfrac{2}{9} = \tfrac{1}{3}$$

$$P(D) = \text{Probability of regions inside } D = \frac{4}{18} + \frac{4}{36} = \frac{1}{3}.$$

$$P(AB) = P(A_1 B_1) + P(A_2 B_2) = \left(\frac{1}{18} + \frac{1}{36}\right) + \left(\frac{1}{9} + \frac{1}{18}\right) = \frac{1}{4}$$

$$P(AC) = \left(\frac{1}{36} + \frac{1}{36}\right) + \left(\frac{1}{18} + \frac{1}{18}\right) = \frac{1}{6}$$

$$P(BC) = \left(\frac{1}{36} + \frac{1}{36}\right) + \left(\frac{1}{18} + \frac{1}{18}\right) = \frac{1}{6}$$

$$P(AD) = P(A_1) = \tfrac{1}{6}; \quad P(BD) = P(B_1) = \tfrac{1}{6}, \quad P(CD) = P(C_1) = \tfrac{1}{9}$$

$$P(ABC) = \frac{1}{36} + \frac{1}{18} = \frac{1}{12}, \quad P(ABD) = P(A_1 B_1) = \frac{1}{12}$$

$$P(BCD) = P(B_1 C_1) = \frac{1}{18}, \quad P(ACD) = P(A_1 C_1) = \frac{1}{18}$$

$$P(ABCD) = P(A_1 B_1 C_1) = \frac{1}{36}.$$

The following results trivially hold

$$P(AB) = P(A)\,P(B),\ P(AC) = P(A)\,P(C),\ P(AD) = P(A)\,P(D)$$

$$P(BC) = P(B)\,P(C),\ P(BD) = P(B)\,P(D),\ P(CD) = P(C)\,P(D)$$

$$P(XYZ) = P(X)\,P(Y)\,P(Z), \text{ all choices}$$

$$P(ABCD) = P(A)\,P(B)\,P(C)\,P(D).$$

Thus, all eleven conditions are satisfied.

***Example* 4.** Prove or disprove :

If events A and B are independent, then $P(AB \mid C) = P(A \mid C) \,.\, P(B \mid C)$.

Alternate Statement. Show that, if A and B are independent events, they need not be independent when conditioned by another event C.

Solution. We throw two regular tetrahedrons and define the following events :

A = {First face odd}, B = {Second face odd}, C = {Total of pips odd}.

We demonstrate the outcome $\omega = \{x, y\}$ and the events as under :

ω	A	B	C	AB	ω	A	B	C	AB
(1, 1)	✓	✓		✓	(3, 1)	✓	✓		✓
(1, 2)	✓		✓		(3, 2)	✓		✓	
(1, 3)	✓	✓		✓	(3, 3)	✓	✓		✓
(1, 4)	✓		✓		(3, 4)	✓		✓	
(2, 1)		✓	✓		(4, 1)		✓	✓	
(2, 2)					(4, 2)				
(2, 3)		✓	✓		(4, 3)		✓	✓	
(2, 4)					(4, 4)				

$$P(A) = \tfrac{8}{16}, P(B) = \tfrac{8}{16}, P(C) = \tfrac{8}{16}, P(AB) = \tfrac{4}{16}$$

$$P(A \mid C) = \tfrac{4}{8}, P(B \mid C) = \tfrac{4}{8}, P(AB \mid C) = 0, \{(x \in AB) \notin C\}$$

Thus, $\quad P(AB) = P(A)\,P(B)$, but $P(AB \mid C) \neq P(A \mid C)\,P(B \mid C)$.

***Example* 5.** If A and B are independent events, show that

$$\max\{P(A \cup B)', P(AB), P(A \,\Delta\, B)\} \geq \tfrac{4}{9}.$$

Solution. Write $P(A) = a$, $P(B) = b$, then we need prove

$P\{(1-a)(1-b), ab, a(1-b) + b(1-a)\} \geq \frac{4}{9}$ ...(1)

Suppose to the contrary that (1) is untrue, then this is

$\max P\{1-(a+b)+ab, ab, a+b-2ab\} < \frac{4}{9}$ ...(2)

By definition of maximum, (2) yields

$$1-(a+b)+ab<\tfrac{4}{9},\quad ab<\tfrac{4}{9},\; a+b-2ab<\tfrac{4}{9} \qquad ...(3)$$

From 3(ii) and 3(iii), $a+b<(\frac{12}{9})=\frac{4}{3}$

So $a+b<(\frac{4}{3})$ and $ab<\frac{4}{9} \;\Rightarrow\; a+(\frac{4}{9}a)<\frac{4}{3}$

$\therefore \quad 9a^2-12a+4<0 \quad\Rightarrow\quad (3a-2)^2<0.$

This is absord, since a is real. Thus we reject (2) and accept (1).

Note that 3(i) and 3(ii) provides $ab<\frac{7}{9}$ different from $ab<\frac{4}{9}$.

Infact regions $a+b<\frac{4}{3}$, $ab<\frac{4}{9}$, $0<a,\, b<1$ have no points, empty region.

***Example* 6.** A player tosses a coin and is to score one point for every head and two for every tail. He is to play on until his score reaches or passes n. If p_n is the chance for attaining exactly n, show that $p_n=\frac{1}{2}(p_{n-1}+p_{n-2})$ and hence find the value of p_n.

Solution. Let S_n denote the score is n and H, T denote head and tail. Then $P(H)=P(T)=\frac{1}{2}$, $P(S_n)=p_n$. Now S_n can be attained by throwing a tail preceded by S_{n-2} or by throwing a head preceded by S_{n-1}. Thus $S_n=(S_{n-1}\cap H)\uplus(S_{n-2}\cap T)$.

$\therefore\; P(S_n)=P(S_{n-1}\cap H)+P(S_{n-2}\cap T)=P(S_{n-1})\,P(H)+P(S_{n-2})\,P(T)$

$$\Rightarrow \quad p_n=\tfrac{1}{2}(p_n+p_{n-2}). \qquad ...(1)$$

To find p_n explicitly, we rewrite (1) as

$$2p_n=p_{n-1}+p_{n-2} \text{ or } 2p_{n+2}=p_{n+1}+p_n.$$

with E as extension (or shift) operator, this is expressible as

$$(2E^2-E-1)\,p_n=0.$$

The roots of the characteristics equation : $2E^2-E-1=0$ are $E=1,-\frac{1}{2}$. Hence

$$p_n=a(1)^n+b(-\tfrac{1}{2})^n=(a+b)\,(-\tfrac{1}{2})^n$$

Now $p_1=P(S_1)=\frac{1}{2}$, $p_2=P(S_2)=\frac{3}{4}$, hence $a=\frac{2}{3}$, $b=\frac{1}{3}$.

$$\therefore \quad p_n=(\tfrac{2}{3})+(\tfrac{1}{3})(-\tfrac{1}{2})^n.$$

Example* 7. *Occupancy Problem. Suppose b distinct balls are randomly distributed among N distinct boxes, write

$$E' = \{\text{no box is empty, i.e. all boxes are occupied}\}$$
$$E_n^{(N)} = \{\text{Exactly } m \text{ boxes remain empty}\}$$

Then,
$$P(E') = \sum_{r=0}^{N}(-1)^r\binom{N}{r}\left(1-\frac{r}{n}\right)^b$$

$$P(E(N)^m) = \binom{N}{m}\left\{\sum_{k=0}^{N-m}(-1)^k\binom{N-m}{k}\left(1-\frac{k-m}{N}\right)^b\right\}.$$

Solution. First Part. First ball has N chocies to land in a box. Infact, all b balls each has N choices to land. Thus b balls can be distributed among N boxes in $(N)^b$ ways. Let E_i denote the event that ith box is empty. Then, we fill the remaining $N - 1$ boxes by b balls, so that $P(E_i) = (N-1)^b/(N)^b$. If two boxes are kept empty, then

$P(E_i\,E_j) = (N-2)^b/(N)^b$, $(i \neq j)$. Similarly, $P(E_i\,E_j\,E_k) = (N-3)^b/(N)^b$, etc.,

Thus, with usual notation

$$S_1 = \Sigma P(E_i) = \binom{N}{1}\frac{(N-1)^b}{(N)^b}, \quad S_2 = \Sigma P(E_i\,E_j) = \binom{N}{2}\frac{(N-2)^b}{(N)^b}, \ldots$$

$$S_r = \sum_{r=1}^{N} P(E_{i_1}\,.\,E_{i_2}\ldots E_{i_r}) = \binom{N}{r}\frac{(N-r)^b}{(N)^b} = \binom{N}{r}\left(i-\frac{r}{N}\right)^b.$$

$P(E') = P(\text{no empty box}) = 1 - P(\text{atleast one empty box}) = 1 - P(E_1 \cup E_2 \cup \ldots \cup E_N)$

$$= 1 - [S_1 - S_2 + S_3 - \ldots + (-1)^N S_N] = \sum_{r=0}^{N}(-1)^r S_r = \sum_{r=0}^{N}(-1)^r\binom{N}{r}\left(1-\frac{r}{N}\right)^b$$

Second Part. For any *specific* group of m empty boxes, we have

$p' = P\{E_1\,E_2\ldots E_m\,\overline{E}_{m+1}\,\overline{E}_{m+2}\ldots\overline{E}_N\}$, [Write $E = E_1, E_2, \ldots, E_m$ and apply Product Rule]

$= P\left\{E\,\overline{E}_{m+1}\,\overline{E}_N\right\} = P(E)\,P\{\overline{E}_{m+1}\ldots\overline{E}_N \mid E\}$. [Now Apply Negation Rule]

$= P(E)\,[1 - P\{(E_{m+1} \cup E_{m+2} \cup \ldots \cup E_N) \mid E\}]$

For smooth notation, write $E_{m+1} = D_1$, $E_{m+2} = D_2$, ... $E_N = D_M$ where $M = N - m$. Now apply Poincare's Rule to get,

$p' = P(E)\ \{1 - (S_1 - S_2 + S_3 - \ldots + (-1)^{r-1} S_r \pm \ldots (-1)^{M-1} S_M\}\quad S_r = \Sigma P\{D_{i_1}, D_{i_2} \ldots D_{i_r} \mid E\}, \ldots\ldots(1)$

Note that $E = m$ boxes empty, $ED_1 = (m+1)$ boxes empty, $ED_1\,D_2 = (m+2)$ boxes empty, ...

$P(E) = (N-m)^b/N^b = [1-(m/N)^b$ [Part 1]

$$P(D_1 \mid E) = \frac{P(ED_1)}{P(E)} = \frac{[N-(m+1)]^b/N^b}{(N-m)^b/N^n} = \left(1-\frac{1}{M}\right)^b, (M = N - m)$$

$$P(D_1\,D_2 \mid E) = \frac{P(ED_1\,D_2)}{P(E)} = \frac{[N-(m+2)]^b/N^b}{(N-m)^b/N^b} = \left(1-\frac{2}{M}\right)^b$$

$$P(D_1\,D_2\ldots D_R \mid E) = \left(1-\frac{k}{M}\right)^b, \ldots$$

Substituting from above results into (1) gives

$$p' = \left(\frac{M}{N}\right)^b\left\{1-\binom{M}{1}.\left(1-\frac{1}{M}\right)^b+\binom{M}{2}\left(1-\frac{2}{M}\right)^b-\binom{M}{3}\left(1-\frac{3}{M}\right)^b+\ldots+(-1)^n\binom{M}{k}\left(1-\frac{k}{M}\right)^b.\pm\ldots\right\}$$

$$= \left(\frac{M}{N}\right)^b\left\{\sum_{k=0}^{M}(-1)^k\binom{M}{k}\left(1-\frac{k}{M}\right)^b\right\}$$

For any group (not a specific group) of m empty boxes, $p = \binom{N}{m} p'$ and using $M = N - m$, the above yields

$$p = \binom{N}{m} \cdot \left(1 - \frac{m}{N}\right)^b \left\{ \sum_{k=0}^{N-m} (-1)^k \binom{N-m}{k} \left(1 - \frac{k}{N-m}\right)^b \right\} \equiv \binom{N}{m} \sum_{k=0}^{N-m} (-1)^k \binom{N-m}{k} \left(1 - \frac{m+k}{N}\right)^b$$

Remarks. $P\{E'\} = P\{m = 0\} = P(\text{No box is empty}\} \sum_{k=0}^{N} (-1)^k \binom{N}{k} \left(1 - \frac{k}{N}\right)^b$

$$P\{\text{Atleast one box is empty}\} = 1 - P(E') \equiv \sum_{k=1}^{N-1} (-1)^{k+1} \binom{N}{k} \left(1 - \frac{k}{N}\right)^b$$

Problems with Solutions Provided at the End of the Text

1*. Given $P(A \cup B) = 5/6$, $P(AB) = 1/3$, $P(\overline{B}) = 1/2$, determine $P(A)$ and $P(B)$. Hence show that events A and B are independent.

2*. Let $P(A_i) = p_i$, $1 \le i \le 3$, $0 < p_i$, 1. Pair-wise independent events A_1, A_2, A_3 satisfy the property : $A_1A_2 = A_2A_3 = A_3A_1 = A_1A_2A_3$. Show that these events are not independent.

3*. Let A, B, C be three events. Show that Ind. (A ; B) and Ind. (B ; C) do not imply Ind. (A ; C). That is, the relation "independent of" is not a transitive relation on the event space (σ-algebra).

4*. Of three events A, B, C suppose that $P(AB) = P(A)\ P(B)$; $P(ABC) = P(A)\ P(B)\ P(C)$. Show that AB and C are independent but A and C need independent.

5*. Prove or disprove the following assertions : if A is indep. of B and A is indep. of C, then (i) A is indep. of $B \cup C$, (ii) A is indep. of $B \uplus C$.

6*. If events A, B, C are indep. with $P(A) = p$, $P(B) = q$, $P(C) = r$, find the probability that atleast two will occur among the three events.

7*. Let A and B be indep. events. Prove or disprove : $A \cup B$ and $A \cap B$ are indep. events.

8*. If the event A is completely independent of events B and C, then A is independent of the logical sum $B \cup C$.

9*. (a) Events A and B are such that $P(A) = \frac{1}{4}$, $P(B \mid A) = \frac{1}{2}$, $P(A \mid B) = \frac{1}{4}$.

Prove or disprove the following statements :

(i) A and B are mutually exclusive. (ii) A is *a* subevent of B.

(iii) $P(\overline{A} \mid \overline{B}) = 3/4$. (iv) $P(A \mid B) + P(A \mid \overline{B}) = 1$.

(b) If A is indep. of B and B is indep. if C and also A is indep. of BC, prove that C is indep. of AB.

10*. For independent events A, B, C let

$P(A) = a$, $P(A \cup B \cup C) = 1 - b$, $P(ABC) = 1 - c$, $P(A'B'\ C) = x$.

Show that the probability x satisfies the condition

$ax^2 + [ab - (1 - a)(a + c - 1)]\ x + b(1 - a)(1 - c) = 0.$ ...(A)

Deduce : $c > [(1 - a)^2 + ab]/(1 - a)$.

11*. A problem is given to three students A, B, C whose chances of solving it are $\frac{1}{2}, \frac{1}{3}, \frac{1}{4}$ respectively. What is the probability that the problem is solved ?

12*. A can hit a target 3 times in 5 shots, B : 2 times in 5 shots, C : 3 times in 4 shots. They fire a volley. Find chance that :

(a) 2 shots hit, (b) At least 2 shots hit, (c) Exactly one of them hits.

If only one of them hits the target, what is the chance that it was the first man ? And if two hit, what is the chance that it is C who missed ?

13*. Events $A_1, A_2, \ldots, A_n$ are independent with $P(A_i) = p_i$. Find the probability p that none of the events occurs and show that $p \le \exp(\Sigma p_i)$.

14*. From the set $S = \{1, 2, \ldots, N\}$, r subsets $A_1, A_2, \ldots, A_r$ are selected in independent trials such that $P\{x_i \in A_i \subset S\} = p_i$, $(q_i = 1 - p_i)$. Find the probability that the subsets A_i, $1 \le i \le r$ do not intersect pairwise.

Exercise 2(b)

1. Assume that none of the events has zero probability. Prove or disprove.

(a) $P(A \mid B) > P(A) \Rightarrow P(B \mid A) > P(B)$ (b) $P(A \mid B) > P(A) \Rightarrow P(A \mid \bar{B}) < P(A)$

(c) $P(A \mid B) \ge P(A) \Rightarrow P(B \mid A) \ge P(B)$ (d) $(A \text{ ind } B) \Rightarrow P(AB \mid C) = P(A \mid C)\, P(B \mid C)$

(e) $P(A \mid B) = P(B) \Rightarrow (A \text{ indep. } B)$ (f) $P(B \mid \bar{A}) = P(B \mid A) \Rightarrow (A \text{ indep } B)$

(g) $P(A) = P(B) \Rightarrow P(A \mid B) = P(B \mid A)$ (h) $P(A \mid B) = P(B \mid A) \Rightarrow P(A) = P(B)$.

2. (i) For any event A, prove that (a) A and ϕ are endep., (b) A and S are indep.

(ii) Suppose A, B, C are independent events. Prove that

(a) $P(B \mid AC) = P(B \mid A \cup C) = P(B)$, (b) A and $(B \,\Delta\, C)$ are indep.

3. (a) Let A and B possible outcomes of an experiment and suppose that

$$P(A) = 0.4, \quad P(B) = p, \quad P(A \cup B) = 0.7.$$

Find p if : (a) A and B are disjoint, (b) A and B are independent.

Find p if : (c) A is subevent of B. (d) $P(A \mid B) = 0.4$. [**Ans.** 0.3, 0.5, 0.7, 0.5]

(b) Find $P(A)$, $P(B)$ and $P(C)$, where mutually Indep. events A, B, C satisfy the relations :

$P(AB) = 1/6$, $P(\bar{B}C) = \frac{1}{2} P(A)$, $3P(B) = 1/P(C)$. [**Ans.** $\frac{1}{3}, \frac{1}{2}, \frac{2}{3}$]

(c) Events $A_1, A_2, \ldots, A_n$ are mutually independent with $P(A_j) = 1 - (1/\alpha)^j$, $1 \le j \le n$. Prove that $P(A_1 \cup A_2 \cup \ldots \cup a_n) = 1 - (\alpha)^{-n(n+1)/2}$.

4. (a) An experiment consists of tossing a coin and a die. If H is the event that head comes up in tossing the coin and D is the event that 3 or 6 comes up in tossing the die, find the probability for the following events : $P(H)$, $P(\bar{D})$, $P(\overline{HD})$, $P(H \mid D)$, $P(\bar{H} \cup \bar{D})$.

[**Ans.** $\frac{1}{2}, \frac{2}{3}, \frac{1}{3}, \frac{1}{2}, \frac{5}{6}$]

(b) A coin is tossed 3 times. Show that the chance of getting head and tail alternately, is $\frac{1}{4}$.

5. (a) The odds against A solving a certain problem are 4 to 3, and odds in favour of B solving the same problem are 7 to 5. Show that the probability that the problem is solved is 16/21. [*Note.* Odds is favour a ; b means $p = a/(a+b)$, $q = b/(a+b)$].

(b) The odds against a certain event are 5 : 2 and odds in favour of another independent event are 6 : 5. Show that the chance that

(i) One at least of the events will happen is $\frac{52}{77}$, (ii) none of the events shall happen is $\frac{25}{77}$.

(c) The odds that a book is favourably reveiwed by three independent critics are 5 : 2, 4 : 3, and 3 : 4 respectively. Show that the chance that of the three critics, a majority is favourable to review, is 209/343.

5. The probability that a teacher will give an unannounced test during any class meeting is $\frac{1}{5}$. If a student is absent twice, show that the prob. that he will miss at least one test is $\frac{9}{25}$.

6. A town has two doctors A and B operating independently. If the probability that doctor A is available is 0.9 and that for B is 0.8, show that the probability that at least one doctor is available when needed, is 0.98.

7. (a) If p is the probability that a man aged x will die in a year, prove that the probability that out of n men $M_1, M_2, \ldots, M_n$ each aged x, M_1 will die in a year and be the first to die, is $(1 - q^n)/n$.

(b) The probability that a husband will live 15 more year is $\frac{1}{4}$, and that his wife will live 15 more-more years is $\frac{1}{3}$. Find the probability that in 15 years hence, (i) both will be alive, (ii) neither will be alive, (iii) at least one will be alive, (iv) only the wife will be alive and (v) only the husband will be alive. **[Ans.** $\frac{1}{12}, \frac{1}{2}, \frac{1}{2}, \frac{1}{4}, \frac{1}{6}$]

8. (*a*) A problem is given to five students A, B, C, D, E. Their chances of solving it are $\frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{4}, \frac{1}{5}$ respectively. Show that the chance that the problem will be solved is 0.85.

(b) A language class has only three students A, B, C and they independently attend the class. The chances of attendance of A, B, C on any given day are $\frac{1}{2}, \frac{1}{3}$ and $\frac{3}{4}$ respectively. Prove that the chance that the total number of attendances in two consecutive days is exactly 3, is $\frac{1}{4}$.

9. (a) Two gunners fire on shot each. The chance of the first gunner getting a hit is 0.7 and that of second getting a hit is 0.6. Show that the chance of at least one scoring a hit, is 0.88.

(b) One shot is fired from each of the three guns. Let G_i denote the event that the target is hit by ith gun, $i = 1, 2, 3$. If $P(G_1) = 0.5$, $P(G_2) = 0.6$, $P(G_3) = 0.8$ and G_i are independent events, prove that the probability that (i) Exactly one hit is registered is 0.26, (ii) At least two hits are registered is 0.70.

(c) When soldiers fire at a target, the chances that they hit the target are : $\frac{1}{3}$ for A, $\frac{1}{6}$ for B, $\frac{1}{6}$ for C, $\frac{1}{12}$ for D. [A, B, C, D are soldiers]. If all the four soldiers fire at the target simultaneously, show that the probability that the target is hit by some one or more is 0.58.

10. (a) The outcome of an experiment is equally likely to be one of the four points in 3-dim. space with coordinates (1, 0, 0), (0, 1, 0), (0, 0, 1) and (1, 1, 1). Let A, B, C be the events : x-coordinate = 1, y-coordinate = 1, z-coordinate = 1 respectively. Show that the events A, B, C are pairwise independent.

(b) An urn contains four tickets marked with number 112, 121, 211, 222 and one ticket is drawn at random. Let A_i ($i = 1, 2, 3$) be the event that ith digit of the no. of the ticket drawn is 1. Prove that the indep. of the events A_1, A_2, A_3, is pairwise.

(c) An urn contains eight tickets marked with numbers 111, 121, 122, 122, 211, 212, 212, 221 and one ticket is drawn at random. Let A_i, ($i = 1, 2, 3$) be the event that ith digit of the

numbers on the ticket drawn is 1. Show that A_i are not pairwise independent although $P(A_1A_2A_3) = P(A_1)\,P(A_2)\,P(A_3)$.

(d) The faces of a regular tetrahedron are coloured as red, green, black and all these colours on the remaining face. Let R, G, B be the events that the face which appears at the bottom, when the tetrahedron is thrown, has the colour red, green, black respectively. Show that the event R, G, B are pairwise independent but are not totally independent.

11. (a) Of three independent events, the probability that the first *only* should happen is $\frac{1}{4}$, the probability that the second *only* should happen is $\frac{1}{8}$ and the probability that the third *only* should happen is $\frac{1}{12}$. Show that the (unconditional) probabilities of the three events are $\frac{1}{2}, \frac{1}{3}, \frac{1}{4}$ respectively.

(b) Of three independent events, the chance that the first *only* should happen is a, the chance of the second *only* is b and the chance of the third *only* is c. Show that the independent chances of the three events are respectively. $a/(a+x)$, $b/(b+x)$, $c/(c+x)$ where x is the root of the equation $(a+x)\,(b+x)\,(c+x) = x^2$.

12. If A and B are independent events such that $AB \subset C$, and $\overline{A}\,\overline{B} \subset \overline{C}$, show that $P(A)\,P(C) \le P(AC)$.

2-30. *Huyghen's Problem*

This is a class of problems whose solutions ultimately depend on an infinite geometric progression. We illustrate the process by a few problems :

***Example* 1.** Two players A and B alternately roll a pair of unbiased dice. A wins if on a throw he obtains 6 before B gets 7 points, B winning in the opposite event. If A starts the game, prove that his chance of winning is 30/61.

Solution. Probability p of getting sum 6 with two dice is 5/36 and sum 7 has prob. 6/36. Denote by $A \mid a$ the event that A starts (takes turn) and A wins. Obviously, A wins on the odd-numbered attempt, thus if suffixes denote trial number, then A wins on :

$$\{A \mid a\} = A_1 \uplus \overline{A}_1\overline{B}_2\,A_3 \uplus \overline{A}_1\overline{B}_2\,\overline{A}_3\overline{B}_4\,A_5 \uplus \ldots \;[\,\overline{A}_i, \overline{B}_j \text{ independent}]$$

Now $\quad P_1\{\overline{A}_1\overline{B}_2\,\overline{A}_3\overline{B}_4 \ldots \overline{B}_{2k}\overline{A}_{2k+1}\} = qq' \cdot qq' \ldots qq' \cdot p = (qq')^k\,p$

where $q = 31/36$ (A does not get sum 6) and $q' = 30/36$ (B does not get sum 7).

$$\therefore \qquad P\{A \mid a\} = \sum_{k=0}^{\infty}(qq')^k\,p = p'\frac{1}{1-qq'} = \frac{5}{36}\cdot\frac{1}{1-(31\times 5/216)} = \frac{30}{61}.$$

***Example* 2.** A, B, C toss a coin in the order mentioned. The first one to throw a head (prob. p) wins. Assuming that the game may continue indefinitely, find their respective chances of winning.

Solution. Let symbol a denote that A starts, then $A \mid a$, $B \mid a$, $C \mid a$ may denote that A wins, B wins, C wins where A started. If suffixes denote trial numbers, then

$$\{A \mid a\} = A_1 \uplus \overline{A}_1\,\overline{B}_2\,\overline{C}_2\,A_4 \uplus \overline{A}_1\,\overline{B}_2\,\overline{C}_3\,\overline{A}_4\,\overline{B}_5\,\overline{C}_6 A_7 \uplus \ldots$$

$$\{B \mid a\} = \overline{A}_1\,B_2 \uplus \overline{A}_1\,\overline{B}_2\overline{C}_3\,\overline{A}_4\,B_8 \uplus \overline{A}_1\,\overline{B}_2\,\overline{C}_3\,\overline{A}_4\,\overline{B}_5\,\overline{C}_6\overline{A}_7 B_8 \uplus \ldots$$

The probability of getting head is p and if tail is $q = 1 - p$.

Now A can win on $(3k+1)$th trial, $0 \le k < \infty$, and B can win on $(3k+2)$th trial, $0 \le k < \infty$.

$P\{\overline{A}_1\overline{B}_2\overline{C}_3\overline{A}_4\overline{B}_5\overline{C}_6 \ldots \overline{C}_{3k}A\} = q^{3k} \cdot p$, $0 \le k < \infty$, [outcomes are independent].

$$\therefore \quad P\{A \mid a\} = \sum_{k=0}^{\infty} pq^{3k} = p.[1/(1-q^3)]$$

Similarly, $P\{\overline{A}_1\overline{B}_2\overline{C}_3 \ldots \overline{C}_{3k}\ A_{3k+1}\overline{A}_{3k+2}\} = q^{3k+1}p \Rightarrow P\{B \mid a\} = \sum_{k=0}^{\infty} pq^{3k} = pq.[1/(1-q^3)]$.

$P\{C \mid a\} = 1 - \{P(A \mid a) + P(B \mid a)\} = 1 - [p/(1-q^3)] - [pq/(1-q^3)] = pq^2/(1-q^3)$.

For the fair coin, $p = 1/2$, and then $P(A \mid a) = 4/7$, $P(B \mid a) = 2/7$, $P(C \mid a) = 1/7$.

Moral. The starter always has an edge over his rivals.

***Example* 3.** A coin is tossed until for the first time the same result appears twice in succession. To every possible outcome requiring n tosses, attribute probability $(1/2)^n$. Find the probability that (i) the experiment ends before the 6th toss, (ii) an even (odd) number of tosses is required for the experiment to end.

Solution. Here $\Omega = \{hh, tt, htt, thh, hthh, thtt, hthtt, \ldots\}$ where [h = head, t = tail]

By hypothesis, $p_n = P\{X = n\} = (1/2)^n$, where X is the number of tosses required.

Observe that : $P(\Omega) = 2\left[\left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^3 + \left(\frac{1}{2}\right)^4 + \ldots\right] = \frac{1}{2}/(1-\frac{1}{2}) = 1.$

(i) $p = P\{X < 6\} = \Sigma P(X = k) = p_2 + p_3 + p_4 + p_5 = 2\left[\left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^3 + \left(\frac{1}{2}\right)^4 + \left(\frac{1}{2}\right)^5\right] = \frac{15}{16}$.

(ii) $P\{X = \text{even}\} = 2\sum_{k=1}^{\infty} P_{2k} = 2\sum_{k=1}^{\infty} (\frac{1}{4})^k = \frac{2}{3}$, $P\{X = \text{odd}\} = \sum_{k=1}^{\infty} P_{2k+1} = \sum_{k=1}^{\infty} (\frac{1}{4})^k = \frac{1}{3}$.

***Example* 4.** A and B throw alternately a pair of unbiased dice, A beginning. A wins if he throws 7 before B throws 6, and B wins if he throws 6 before A throws 7. If $A \mid a$ and $B \mid a$ denote the events that A wins and B wins when A starts, etc. show that :

(i) $P(A \mid a) = \frac{1}{6} + \frac{5}{6} P(A \mid b)$. (ii) $P(A \mid b) = \frac{31}{36} P(A \mid a)$

(iii) $P(B \mid a) = \frac{5}{6} P(B \mid b)$ (iv) $P(B \mid b) = \frac{5}{36} + \frac{31}{36} P(B \mid a)$.

Hence or otherwise find $P(A \mid a)$ and $P(B \mid a)$. Also comment on the result :

$$P(A \mid a) + P(B \mid a) = 1.$$

[***Note.*** This problem outlines the recurrence relations as an alternative to an infinite G.P.].

Solution. For 2-dice sum, $P(S_k) = \{6 - | 7 - k |\}$, so $P(S_6) = 5/36$, $P(S_7) = 6/36 = 1/6$.

(i) *By Multistage p-Rule :* $A_a = \{A \text{ wins when } A \text{ starts}\}$, etc., $S_A = A$'s winning score, etc.

$$P(A_a) = P(S_A)\, P\{A_a \mid S_A\} + P(\overline{S}_A)\, P\{A_a \mid \overline{S}_A\}, [S_A = S_7, S_B = S_6] \qquad \ldots(1)$$

$P\{A_a \mid S_A\} = 1$, $P\{A_a \mid \overline{S}_A\} = P\{A_b\}$, because $P\{A_a \mid \overline{S}_A\}$ is the probability that wins but after the first trial. But at trial 2, A becomes the second player. Hence (1) yields

$$P(A_a) = (1/6) \cdot 1 + (5/6)\, P\{A_b\}.$$

(ii) $$P(A_b) = P(S_B)\,P(A_b \mid S_B) + P(\bar{S}_B)\,P(A_b \mid \bar{S}_B)$$

$P(A_b \mid S_B) = P\{A$ wins at B's turn when S_B happens$\} = 0,\ P(A_b \mid \bar{S}_B) = P(A_a),\ P(\bar{S}_B) = (\frac{31}{36})$

$\therefore$ $$P(A_b) = (\tfrac{31}{36})\,P(A_a).$$

(iii) $$P(B_a) = P(S_A)\,P(B_a \mid S_A) + P(\bar{S}_A)\,P(B_a \mid \bar{S}_A)$$

Now $$P(B_a \mid S_A) = 0,\quad P(B_a \mid \bar{S}_A) = P(B_b),\ P(\bar{S}_A) = \tfrac{5}{6}$$

$\therefore$ $$P(B_a) = (5/6)\,P(B_b).$$

(iv) $$P(B_b) = P(S_B)\,P(B_b \mid S_B) + P(\bar{S}_B)\,P(B_b \mid \bar{S}_B)$$

Now $$P\{B_b \mid S_B\} = 1,\ P\{B_b \mid \bar{S}_B\} = P(B_a),\ P(S_B) = P(S_6) = \tfrac{5}{36}$$

$\therefore$ $$P(B_b) = (\tfrac{5}{36}) + (\tfrac{31}{36})\,P(B_a)$$

From (i) and (ii), $P(A_a) = \dfrac{1}{6} + \dfrac{5}{6}\cdot\dfrac{31}{36}P(A_a) \Rightarrow P(A_a) = \dfrac{36}{61}$

From (ii) and (iv), $P(B_a) = \dfrac{5}{6}\left\{\dfrac{5}{36} + \dfrac{31}{36}P(B_a)\right\} \Rightarrow P(B_a) = \dfrac{25}{61}$.

Obviously, $P(A_a) + P(B_a) = 1$, which is otherwise obtainable by Negation Rule.

Exercise 2(c)

1. A and B take alternate turns in throwing a coin, the first to throw head being awarded a prize. Show that if A has the first throw, the chances of their winning are in the ration 2 : 1. It is assumed that the game is continued till one of the players wins it.
2. A and B alternate shots in a dual. The probability of A hitting B is p_1 and probability of B hitting A is p_2. If A fires first, show that the chance that A wins the dual is $p_1/(1 - q_1 q_2)$.
3. An urn contains a white balls and b black balls. Two players draw balls out of it in turn, one ball at a time and each time return the ball to the urn. Show that the chance that the player who starts will first get a white ball is $(a + b)/(a + 2b)$.
4. A and B take turns in throwing 2 dice, the first to throw 9 being awarded the prize. Show that if A starts, their chances of winning are in the ratio 9 : 8.
5. A man alternately tosses a coin and throws a die, beginning with a coin. Show that the chance that he will get a head before he gets a "5 or 6" on the die is 3/5.
6. A, B, C, D cut a peak of cards successively in the order mentioned. What are their respective chances of first cutting a spade ? **[Ans.** $64\lambda, 48\lambda, 36\lambda, 27\lambda, \lambda = 1175$]
7. A, B, C alternate firing at a target in this order. The probabilities that A, B, C strike the target are p_1, p_2, p_3 respectively. Assuming Bernoulli trials, show that the chance that A will hit the target first is $p_1/(1 - q_1 q_2 q_3)$ and the chance that B will hit the target first is
$$q_1 p_2/(1 - q_1 q_2 q_3)\ .\ [q = 1 - p].$$
8. N players $A_1, \ldots, A_N$ throw in the order mentioned, a biased die whose probability of a six is p. The first one to throw a six wins. Show that $P\{A_k \text{ wins}\} = pq^{k-1}/(1 - q^N),\ 1 \le k \le N$.

2-40. *Successive Trials*

If an event A occurs at random on an average once in time t, the chance that A does not occur in given time T is e^T/t.

Proof. Divide every unit of time in n intervals, n being very large. Then in the periods t and T, there shall be nt and nT such intervals. Consequently, $P(A) = 1/nt$, since A happens in one interval out of nt intervals.

Now, for a single interval $P(\overline{A}) = 1 - (1/nt) = q$. If A does not occur in the entire set of nT intervals, then its probability is $q \cdot q \ldots q$ [these are nT terms] $= (q)^{nT}$. Hence,

$$P\{A \text{ does not occur in time } T\} = \lim_{n\to\infty}\left(1-\frac{1}{nt}\right)^{nT} = e^{-T/t}. \qquad \text{[by Euler's limit]}$$

***Example* 1.** If war breaks out on the average, once in 25 years, find the probability that in 50 years at a stretch, there shall be no war.

Solution. Divide the period of 25 years into an indefinitely large number, say, n equal intervals. Then, the probability of a war in any of these equal intervals, assumed equi-probable, is $p = (1/n)$, whence that of no war breaking out is $q = 1 - (1/n)$. Now the period of 50 years gives rise to $50 \times (n/25) = 2n$, equi-probable sub-periods. Hence the probabiliy that there will be no war in the next 50 years is $(q \cdot q \ldots. q)$ which are $2n$ terms, i.e., q^{2n}. Thus for large n,

$$P = \lim_{n\to\infty}(q^{2n}) = \lim_{n\to\infty}\left(1-\frac{1}{n}\right)^{2n} = e^{-2}. \qquad \text{[by Euler's limit]}$$

***Example* 2.** If 30 D.T.C. buses per hour pass University Campus, what is the chance that in the next 10 minutes none will pass ?

Solution. Here $t = 60/30$, $T = 10$, hence $P\{\text{No bus passes}\} = e^{-10/2} = e^{-5} = 0.00674$.

2-50. *Baye's Reversal Rule*

If $A_1, A_2, \ldots, A_n$ are mutually disjoint events, with $P(A_i) \neq 0$, $1 \leq i \leq n$, and if B is any arbitrary event caused by events A_j, so $B \subseteq \uplus_i A_i$, with $P(B) > 0$, then

$$P(A_i \mid B) = \frac{P(A_i)\,P(B \mid A_i)}{P(B)} = \frac{P(A_i)\,P(B \mid A_i)}{\Sigma P(A_i)\,P(B \mid A_i)}. \qquad \ldots(1)$$

Proof. Since $B \subseteq \uplus_i A_i$ it follows that [See Figure §2-10], $B = B \cap (\uplus A_i) = \uplus BA_i$ by distributive law. Taking probabilities of both sides, using finite-additivity [$P3'$], we get

$$P(B) = P(\uplus BA_i) = \Sigma_p(Ba_i) = \Sigma_p(A_i)\,P(B \mid A_i), \quad 1 \leq i \leq n \qquad \ldots(i)$$

Now, by definition of conditional probability $P[A_i \mid B] = \dfrac{P(A_i \cap B)}{P(B)} = \dfrac{P(A_i)\,P(B \mid A_i)}{P(B)}$ ...(ii)

Substituting for $P(B)$ from (i) into Den. of (ii) we obtain the result

$$P\{A_i \mid B\} = \frac{P(A_i)\,P\{B \mid A_i\}}{\Sigma P(A_i)\,P(B \mid A_i)}, \quad 1 \leq i \leq n$$

Note. Baye's Reversal Rule is also termed "Baye's Retrodiction Formula".

2-51. *Baye's Theorem for Future Events*

If $A_1, A_2, ..., A_n$ are mutually disjoint events, with $P(A_i) \neq 0$, $i = 1, 2, ..., n$ and if B is any event, $[P(B) > 0, B \subseteq \uplus A_i]$ which is, conditioned on the occurrence of event C, $[P(C > 0, C \subseteq \uplus A_i]$, then

$$P\{B \mid C\} = \frac{\Sigma P(A_i)\, P(C \mid A_i)\,.\, P(B \mid CA_i)}{\Sigma P(A_i)\, P(C \mid A_i)}, 1 \le i \le n. \qquad ...(1)$$

Proof. Since $B \in \uplus A_i$, we get, by Distributive law, $B = B \cap \uplus A_i) = \uplus(BA_i)$, $1 \le i \le n$

$\therefore \quad P\{B \mid C\} = P\{\uplus BA_i \mid C\} = \Sigma P(BA_i \mid C)$ ...(i)

where we accept that $P\{* \mid C\}$ is *a* bonafide probability measure and use finite additivity [$P3'$]. Using the definition of conditional probability, we have

$P\{BA_i \mid C\} = P(A_iBC)/P(C) = P(A_i)\,.\,P(C \mid A_i)\,.\,P(B \mid CA_i)/P(C)$, [by Product Rule] ...(ii)

Since $C \subseteq \uplus A_i$, so $C = C \cap (\uplus A_i) = \uplus(CA_i)$, hence $P(C) = \Sigma P(CA_i) = \Sigma P(A_i)\, P(C \mid A_i)$ ...(iii)

with obvious use of Distributive law, finite-additivity axiom [$P3'$] and use of Product Rule. Substituting from (ii) and (iii) into (i), the result (1) follows.

Cor. Suppose occurrence of A_i makes B independent of C then $P(BC \mid A_i) = P(B \mid A_i)\, P(C \mid A_i)$.

$\therefore \qquad P(BCA_i) = P(A_i)\, P(BC \mid A_i) = P(A_i)\,.\,P(B \mid A_i)\,.\,P(C \mid A_i)$ [Conditional indep.]

Also $\qquad P(BCA_i) = P(CA_i)\, P(B \mid CA_i) = P(A_i)\,.\,P(C \mid A_i)\, P(B\, CA_i)$

These provide : $P(B \mid CA_i) = P(B \mid A_i)$; and using this result in (1) yields.

$$P\{B \mid C\} = \frac{\Sigma P(A_i)\, P(C \mid A_i)\,.\, P(B \mid A_i)}{\Sigma P(A_i)\, P(C \mid A_i)}, 1 \le i \le n \qquad ...(2)$$

We may regard event B as a future event in relation to event C.

2-52. *Baye's Theorem for Repeated Trials*

In n independent trials with constant probability p of success in each trial, r successes are observed. Probability p is not known but it is known that it takes one of the values $p_1, p_2, ..., p_k$ with $P(p = p_i) = \alpha_i$, $i = 1, 2, ..., k$. Investigate $P\{\alpha \le p \le \beta\}$, where $0 \le \alpha \le \beta \le 1$.

Proof. We designate the evens as under :

$$A = \{r \text{ successes in } n \text{ trials}], B_i = \{p = p_i\}, C = \{p : \alpha \le p \le \beta\}, \cup B_i = S$$

Using Multi-Stage Rule, and noting $P(B_i) = \alpha_i$, $i \le i \le k$, we get

$$P(A) = \Sigma P(B_i)\, P(A \mid B_i) = \sum_{i=1}^{k} \alpha_i \binom{n}{r} p_i^r q_i^{n-r} = \binom{n}{r} \Sigma \alpha_i p_i^r q_i^{n-r} \qquad ...(i)$$

Using Bayes Reversal Rule, we get

$$P(B_i \mid A) = \frac{P(A \mid B_i)\, P(B_i)}{P(A)} = \frac{\binom{n}{r} p_i^r q_i^{n-r}\,.\alpha_i}{\binom{n}{r} \Sigma p_i^r q_i^{n-r}\,.\alpha_i} = \frac{p_i^r q_i^{n-r} \alpha_i}{\Sigma \alpha_i p_i^r q_i^{n-r}}, \qquad \text{[by (i)]} \qquad ...(ii)$$

$$\therefore \quad P(C) = \sum_{p_i=\alpha}^{\beta} P(B_i \mid A) = \sum_{p_i=\alpha}^{\beta} p_i' q_i^{n-i} \alpha_i \Big/ \sum_{i=1}^{k} p_i' q_i^{n-i} \alpha_i. \qquad [p_i + q_i = 1]$$

2-53. Worked-out Problems

***Example* 1.** A writes a letter to B and does not receive a reply. Assuming that one letter in n is lost in the mails, find the chance that B received the letter. It is certain that B would have replied the letter if he had received it.

Solution. Define events : $N = \{A \text{ did not receive reply from } B\}$, $R = \{B \text{ received the letter}\}$. We need to find $P\{R \mid N\}$, where

$$P\{N \mid \bar{R}\} = 1,\ P\{N \mid R\} = 1/n,\ P\{\bar{R}\} = 1/n \text{ (loss in mails)},\ P\{R\} = 1 - P\{\bar{R}\}.$$

Now $\qquad P(N) = P(R)\,P(N \mid R) + P(\bar{R})\,P(N \mid \bar{R})$

$$= (1 - n^{-1})(1/n) + (1/n)\,.\,1 = (2n-1)/n^2.$$

$\therefore\ P\{R \mid N\} = P\{N \mid R\},\ P(R)/P(N) = (n-1)/(2n-1).$ $\qquad$ [Reversal Identity]

***Example* 2.** Urns A, B, C have the following coloured balls : A : red, 4 white, $B = 2$ red, 6 white; C : 1 red, 8 white. (i) An urn is chosen at random; a ball drawn turns out to be red. Find the chance that urn A is chosen.

(ii) An urn is chosen at random; two balls withdrawn without replacement; both turn out to be red. Find the chance that urn A or B or C is chosen.

Solution. Since urn-selection is equiprobable, $P(A) = P(B) = P(C) = 1/3$. Let R_j indicate the appearance of j red balls. Then.

$$P(R_1 \mid A) = 6/10,\ P(R_1 \mid B) = 2/8,\ P(R_1 \mid C) = 1/9.$$

$$P(R_2 \mid A) = \binom{6}{2}\binom{10}{2} = \tfrac{1}{3},\ P(R_2 \mid B) = \binom{2}{2}\binom{8}{2} = \tfrac{1}{28},\ P(R_2 \mid C) = 0.$$

$P(R_1) = P(A)\,P(R_1 \mid A) + P(B)\,P(R_1 \mid B) + P(C)\,P(R_1 \mid C) = 173/540.$ [By Multi-Stage p-Rule]

$P(R_2) = P(A)\,P(R_2 \mid A) + P(B)\,P(R_2 \mid B) + P(C)\,P(R_2 \mid C) = 31/252.$ $\quad$ By Baye's Reversal Rule

$$\text{(i)}\ P(A \mid R_1) = \frac{P(A)\,P(R_1 \mid A)}{P(R_1)} = \frac{1}{3}\cdot\frac{6}{10} = \frac{173}{540} = \frac{108}{173}.$$

$$\text{(ii)}\ P(A \mid R_2) = \frac{P(A)\,P(R_2 \mid A)}{P(R_2)} = \frac{1/9}{31/252} = \frac{28}{31},\ P(B \mid R_2) = \frac{P(R_2 \mid B)\,P(B)}{P(R_2)} = \frac{1/84}{31/252} = \frac{21}{217}.$$

$$P(C \mid R_2) = P(R_2 \mid C)\,P(C)/P(R_3) = 0.$$

***Example* 3.** The contents of urns I, II, III are 1 white, 2 black, 3 red balls; 2 white, 1 black, 1 red ball; 4 white, 5 black, 3 red balls respectively. One urn is chosen at random and 2 balls drawn. They happen to be white and red. What is the probability that they come from urns, I, II or III ?

Solution. Let $U_1 = \{\text{urn } i \text{ is chosen}\}$, $A = \{\text{one white and one red ball chosen}\}$.

Here $P(U_1) = P(U_2) = P(U_3) = \frac{1}{3}$

$$P(A \mid U_1) = 1 \times 3/\binom{6}{2} = 1/5,\ P(A \mid U_2) = 2 \times 1/\binom{4}{2} = 1/3\ \ P(A \mid U_3) = 4 \times 3/\binom{12}{2} = 2/11.$$

$\therefore\ P(A) = P(U_1)\, P(A \mid U_1) + P(U_2)\, P(A \mid U_2) + P(U_3)\, P(A \mid U_3) = 118/495.$ [Multi-stage p-Rule]

We now utilize Baye's Reversal Rule to obtain

$$P(U_1 \mid A) = \frac{P(U_1)\, P(A \mid U_1)}{P(A)} = \frac{1/15}{118/495} = \frac{33}{118};\ P(U_2 \mid A) = \frac{1/9}{118/495} = \frac{55}{118};\ P(U_3 \mid A) = \frac{2/33}{118/495} = \frac{30}{118}.$$

Note. $P(U_3 \mid A) = 1 - [P(U_1 \mid A) + P(U_2 \mid A)]$.

***Example* 4.** Guns 1 and 2 are shooting at the same target. Gun 1 shoots on the average nine shots during the same time gun 2 shoots ten shots. The precision of these two guns is not the same. On the average, out of 10 shots from gun 1, eight hit the target, and from gun 2, only seven.

In the course of shooting, the target has been hit by a bullet, but it not known which gun shot this bullet. Find the chance that the target was hit by gun 2.

Solution. Let G_i = {a bullet is shot by gun i}, $i = 1, 2$, and H = {target is hit by a bullet}. Taking into account the ratio of the average number of shots made by gun 1 to the average number of shots made by gun 2, we can write $P(G_1) = (9/10)$, $P(G_2) = (9/10)p$, say.

From the data relating to the precision of the guns, we have $P(H \mid G) = 0.8$, $P(H \mid G_2) = 0.7$.

$\therefore\ P(H) = P(G_1)\, P(H \mid G_1) + P(G_2)\, P(H \mid G_2) = (0.9)p \times (0.8) + p(0.7) = (1.42)p$ [Multi-stage p-Rule]

Baye's Reversal Law now gives

$$P(G_2 \mid H) = \frac{P(G_2)\, P(H \mid G_2)}{P(H)} = \frac{(0.7)p}{(1.42)p} = \frac{35}{71} = 0.493.$$

***Example* 5.** In a bolt factory, machines, A, B, C manufacture respectively 25%, 35%, 40% of the total. Of their output, 5%, 4%, 2% are defective bolts. A bolt is drawn at random from the product and is found to be defective. What are the probabilities that it was manufactured by machines A, B and C ?

Solution. Let A, B, C also denote the respective events that a bolt selected a random is manufactured by the machines A, B, C. Let D denote the event that a bolt is 'defective'. Then

$P(A) = 0.25$, $P(B) = 0.35$, $P(C) = 0.40$; $P(D \mid A) = 0.05$, $P(D \mid B) = 0.04$, $P(D \mid C) = 0.02$

$P(D) = P(A)\, P(D \mid A) + P(B)\, P(D \mid B) + P(C)\, P(D \mid C)$ [Multi-stage p-Rule]

$= (0.25)\,(0.05) + (0.35)\,(0.04) + (0.40)\,(0.02) = 345/(10)^4 = 0.0345.$

The posteriori probabilities, (i.e. after information that item is defective) are obtainable from Baye's Reversal Law :

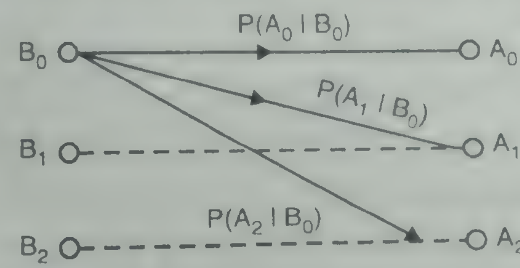
$P(A \mid D) = P(A)\, P(D \mid A)/P(D) = (0.25)\,(0.05)/(0.0345) = 125/345.$

$P(B \mid D) = P(B)\, P(D \mid B)/P(D) = (0.04)\,(0.35)/(0.0345) = 140/345.$

$P(C \mid D) = P(C)\, P(D \mid C)/P(D) = (0.40)\,(0.02)/(0.0345) = 80/345.$

Note. $P(C \mid D) = 1 - [P(A \mid D) + P(B \mid D)]$.

***Example* 6.** A communication system transmits symbols 0, 1, 2 over a channel to a receiver. The channel often causes errors. Let B_i and A_j denote events that the symbols **Before** channel and **After** channel are i and j respectively, $(i, j = 0, 1, 2)$. Assume transmission apiori probabilities are $P(B_0) = 0.5$, $P(B_1) = 0.3$ and $P(B_2) = 0.2$ and transition probabilities are $P(A_i \mid B_j) = 0.1$, $i \neq j$ and $P(A_i \mid B_j) = 0.8$, $i = j$. Compute the reception probabilities $P(A_i)$ and the posteriori probabilities $P(B_j \mid A_i)$ for this system. What happens if all transmission probabilities are equal ?



Solution. (i) The given data : $P(B_0) = 0.5$, $P(B_1) = 0.3$, $P(B_2) = 0.2$

$$P(A_0 \mid B_0) = P(A_1 \mid B_1) = P(A_2 \mid B_2) = 0.8$$

$$P(A_0 \mid B_1) = P(A_0 \mid B_2) = P(A_1 \mid B_0) = P(A_1 \mid B_2) = P(A_2 \mid B_0) = P(A_2 \mid B_1) = 0.1.$$

$$\therefore \quad P(A_0) = P(B_0)P(A_0 \mid B_0) + P(B_1)P(A_0 \mid B_1) + P(B_2)P(A_0 \mid B_2) = (0.5)(0.8) + (0.3)(0.1) + (0.2)(0.1) = 0.45.$$

$$P(A_1) = (0.5)(0.1) + (0.3)(0.8) + (0.2)(0.1) = 0.31, \; P(A_2) = (0.5)(0.1) + (0.3)(0.1) + (0.2)(0.8) = 0.24.$$

$$P(B_0 \mid A_0) = \frac{P(B_0)\,P(A_0 \mid B_0)}{P(A_0)} = \frac{(0.5)(0.8)}{0.45} = \frac{8}{9} = 0.889. \qquad \text{[Baye's Reversal Rule]}$$

$P(B_0 \mid A_1) = (0.5)\ (0.1)/(0.31) = 5/31$; $P(B_0 \mid A_2) = (0.5)\ (0.1)/(0.24) = 5/24$

$P(B_1 \mid A_0) = (0.3)\ (0.1)/(0.45) = 1/15$; $P(B_1 \mid A_1) = (0.3)\ (0.8)/(.31) = 24/31$.

$P(B_1 \mid A_2) = (0.3)\ (0.1)/(0.24) = 1/8$; $P(B_2 \mid A_0) = (0.2)\ (0.1)/(0.45) = 2/45$

$P(B_2 \mid A_1) = (0.2)\ (0.1)/(0.31) = 2/31$; $P(B_2 \mid A_2) = (0.2)\ (0.8)/(0.24) = 2/3$.

(ii) When all transmission probabilities are equal $P(B_i) = 1/3$, $i = 0, 1, 2$. So

$$P(A_0) = (1/3)\ [P(A_0 \mid B_0) + P(A_0 \mid B_1) + P(A_0 \mid B_2)] = (1/3)\ [0.8 + 0.1 + 0.2] = 1/3.$$

Similarly $P(A_1) = 1/3$ and $P(A_2) = 1/3$. This should be obvious in channel transmission.

Now $\quad P\{B_i \mid A_j\} = P(B_i)\ P\{A_j \mid B_i\}/P(A_j) = P\{A_j \mid B_i\}$

Thus, $\quad P\{B_i \mid A_j\} = P\{A_i \mid B_j\} = 0.8,\ i = j;\ P\{B_i \mid A_j\} = 0.1,\ i \neq j$.

***Example* 7.** In answering a question on a multiple choice test, an examinee either knows the answer (with probability p) or he guesses (with probability $q = 1 - p$). Let the probability of answering the question correctly be 1 for an examinee whose knows the answer and $1/m$ for one who guesses (m being the number of multiple choice alternatives). Suppose an examinee answers a question correctly, what is the probability p_m that he really knows the answer ?

Note. $P_m > p$ and sequence $<p_m>$ is strictly increasing for $m \geq 2$.

$$\text{Now } p_m - p = \frac{mp}{mp+q} = \frac{(m-1)pq}{mp+q} > 0, \text{ hence } p_m > p.$$

$$p_{m+1} - p_m = \frac{pq}{[(m+1)p+q][mp+q]} > 0 \Rightarrow p_{m+1} > p_m \text{ so } <p_m> \uparrow.$$

Solution. Let K, G, C denote the respective events that the examinee 'knows the answer', 'guesses the answer' and 'answer correctly'. Then,

$$P(K) = p,\ P(G) = q,\ P\{C \mid K\} = 1,\ P\{C \mid G\} = 1/m.$$

$$\therefore\ P(C) = P(K)\,P(C \mid K) + P(G)\,P(C \mid G) = p\,.\,1 + q(1/m) = (mp + q)/m. \quad \text{[Multi-Stage } p\text{-Rule]}$$

Thus, $P(K \mid C) = P(C \mid K)\,P(K)/P(C) = mp/(mp + q)$ [Baye's Reversal Law]

***Example* 8.** A box having n balls, was filled up in the following manner : An unbiased coin was tossed n times and according as it showed head or tail, one white or one black ball was put into the box. We draw m balls from this box, one by one with replacement, and observe that every one turns out to be white. Find $p_{n,\,m} = P\{\text{Box contains only white balls}\}$. Deduce that

$$p_{n,\,m} > [1 + e^{-\theta}]^{-n} > (1/2)^n\ .\ (\theta = m/n)$$

Solution. Write $W_j = \{\text{Box contains } j \text{ white balls, hence } (n - i) \text{ black ones}\}$. This amounts to j heads and $(n - j)$ tails of n throws of the coin, hence

$$P(W_j) = \binom{n}{j}\left(\tfrac{1}{2}\right)^{n-j}\left(\tfrac{1}{2}\right)^{j} = \binom{n}{j}\left(\tfrac{1}{2}\right)^{n},\ j = 1, \ldots, n \qquad \text{(Warning : } j \neq 0) \quad \ldots(1)$$

Let $A = \{m \text{ balls drawn from the box are all white}\}$. In a box with j white balls among n balls, the probab of a white-ball extraction is (j/n) and for m white extractions in repeated independent trials, this probability is $(j/n)^m$. Hence

$$P\{A \mid W_j\} = (j/n)^m,\ j = 1, 2, \ldots, n$$

$$\therefore\quad P(A) = \sum_{j=1}^{n} P(W_j)\,P(A \mid W_j) = \sum_{1}^{n}\binom{n}{j}\left(\frac{1}{2}\right)^{n}\left(\frac{j}{n}\right)^{m} = \left(\frac{1}{2}\right)^{n}\sum_{1}^{n}\binom{n}{j}\left(\frac{j}{m}\right)^{m}$$

So $\quad P(W_n \mid A) = P(A \mid W_n)\,.\,P(W_n)/P(A)$ [Reversal Identity]

Now $\quad P(A \mid W_n) = 1$ (certainty) $\qquad P(W_n) = (1/2)^n$. [from (1)]

$$\therefore\quad P(W_n \mid A) = \left\{\sum_{1}^{n}\binom{n}{3}\left(\frac{j}{n}\right)^{m}\right\}^{-1}$$

This is the required probability.

Approximation. For $x > 0$, $(1 - x) < e^{-x} \Rightarrow [1 - (j/n)] < e^{-j/m}$. Now,

$$\sum_{j=1}^{n}\binom{n}{j}\left(\frac{j}{n}\right)^{m} = \sum_{j=0}^{n}\binom{n}{j}\left(1 - \frac{j}{n}\right)^{m} < \sum_{0}^{n}\binom{n}{j}(e^{-j/n})^{m} = \sum_{j=0}^{n}\binom{n}{j}(e^{-\theta})^{j} = (1 + e^{-\theta})^{n}$$

$$\therefore\quad p(n, m) \geq (1 + e^{-\theta})^{-n} > (1/2)^n$$

Since $(1/2)^n$ is the aprori probab. that the box contains only white balls. Note that $p(n, m) \to 1$ as $m \to \infty$.

***Example* 9.** An urn contains N balls of which D are defective. A sample of n balls is drawn from the urn. Let A_k be the event that the sample contains exactly k defective balls and B_j the event that the jth ball is defective. Find $(B_j \mid A_k)$ when the sample is drawn (i) without replacement and (ii) with replacement.

Solution. (i) $P(A_k) = \binom{D}{k}\binom{N-D}{n-k}\Big/\binom{N}{n}$; $P(A_k \mid B_j) = \binom{D-1}{k-1}\binom{N-D}{n-k}\Big/\binom{N-1}{n-1}$.

Let E_i denote the event that exactly i defectives occur in the first $(j-1)$ draws, then

$$P(E_i) = \binom{D}{i}\binom{N-D}{j-1-i}\Big/\binom{N}{j-1};\ P(B_j \mid E_i) = \frac{D-i}{N-(j-1)}.$$

$$\therefore\ P(B_j) = \sum_{i=0}^{j-1} P(B_j \mid E_i)\,P(E_i) = \sum_{i=0}^{j-1}\frac{D-i}{N-j+1}\binom{D}{i}\binom{N-D}{j-1-i}\Big/\binom{N}{j-1} = \frac{D}{N}.\ \text{[Multi-stage Rule]}$$

So, $\quad P(B_j \mid A_k) = P(A_k \mid B_j)\,P(B_j)/(A_k).$ [Baye's Reversal Rule]

$$= \frac{\left[\binom{D-1}{k-1}\binom{N-D}{n-k}\Big/\binom{N-1}{n-1}\right]\cdot\frac{D}{N}}{\binom{D}{k}\binom{N-D}{n-k}\Big/\binom{N}{n}} = \frac{k}{n}.$$

(ii) $$P(A_k) = \binom{n}{k}\frac{D^k(N-D)^{n-k}}{(N)^n};\ P(A_k \mid B_i) = \binom{n-1}{k-1}\frac{D^{k-1}(N-D)^{n-k}}{(N)^{n-1}}$$

Since the balls are repalced after each draw, $P(B_j) = D/N$, $\forall\, j$. Hence, by Baye's Reversal Rule,

$$P(B_j \mid A_k) = \frac{P(A_k \mid B_j)\,P(B_j)}{P(A_k)} = \frac{\left[\binom{n-1}{k-1}\frac{D^{k-1}(N-D)^{n-k}}{(N)^{n-1}}\right]\frac{D}{N}}{\binom{n}{k}\frac{D^k(N-D)^{n-k}}{(N)^n}} = \frac{k}{n}.$$

Observe that the same result obtains in either scheme.

Problems with Solutions Provided at the End of the Text

1*. If a toss of three dice gives a total of 9, what is the probability that this total was produced by three 3's ?

2*. Coin C_1 is unbiased and coin C_2 is 2-headed and $P\{C_1$ is chosen$\} = 3/4$. A coin is chosen at random and tossed once. Find $P\{C_2$ is chosen $\mid$ head had appeared$\}$.

3*. The chances of A, B, C becoming managers of a certain company are 4 : 2 : 3. The probabilities that Bonus Scheme shall be introduced if A, B, C become managers are 0.3, 0.5, 0.8 respectively. Find the probability that the Bonus Scheme will be introduced.

If the Bonus Scheme has been introduced, what is the probability that A is appointed as the manager ?

4*. A bag contains an assortment of blue and red balls. If two balls are drawn at random, the probability of drawing of 2 red balls is five times the probability of drawing 2 blue balls. Further more, the probability of drawing 1 ball of each colour is six times the probability of drawing 2 blue balls. Find the number of red and blue balls in the bag.

5*. Three urns A_1, A_2, A_3 contain respectively 3 red, 4 white, 1 blue ; 1 red, 2 white, 3 blue; 4 red, 3 white, 2 blue balls. One urn is chosen at random and a ball is withdrawn. It is found to be red. Find the probability that it came from urn A_2.

6*. In a certain college, the geographical distribution of men students is as follows : 50% come from the East, 30% come from the Mid-West, 20% come from the Far-West. The following proportions of the men students wear ties : 80% of the Easterners, 60% of the Mid-Westerners, and 40% of the Far-Westerners. What is the prob. that a student who wears a tie comes from the East ? From the Mid West ? From the Far-West ?

7*. Among three identical chests, one contains 2 black balls, one contains 1 black and 1 red ball and one contains 2 red balls. A chest is chosen at random and a ball is taken out. If it is red, what is the prob. that the other ball in the chest is also red ?

8*. *A* tells the truth with probability p and *B* with probability p'.

(i) If they make the same statement, what is the chance that the statement is true.

(ii) If *A* makes a statement and *B* says that *A* lies, what is the chance that *A* told the truth ?

9*. Of three identical boxes *A*, *B*, *C*, box *A* contains 1 white, 2 black chips, box *B* contains 2 white, 1 black chip and box *C* contains 2 white, 2 black chips. One of the boxes is chosen at random and one chips is drawn. It turns out to be white. Find the chance of drawing a white chip again, without replacements.

10*. Box *A* contains 4 red, 2 white and 6 black balls and box *B* contains 3 red and 5 white balls. *A* fair die is tossed. If 1 or 6 appears, a ball is chosen from *A*, otherwise a ball is chosen from *B*. If a red ball is chosen, what is the chance that a '6' appeared on the die ?

11*. Urn *A* contains 2 white and 2 black balls. Urn *B* contains 3 white and 2 black balls. One ball is transferred from *A* to *B*, and then one ball is drawn out of *B*. Find the chance that this ball is white. If this ball turns out to be white, find the probability that the transferred ball was white.

12*. A committee of $6n$ people contains half men and half women. A male is elected to choose, at random. From the remaining $6n - 1$ people, a committee of n people is formed. If the n people chosen happen to be of the same sex, find the probability that they are all female.

13*. From a vessel containing 3 white and 5 black balls, 4 balls are transferred into an empty vessel. From this vessel, a ball is drawn and it is found to be white. What is the chance that out of 4 balls transferred 3 are white and 1 is balck ?

14˙. The probability that a family has k children is P_k, $(1 \le k \le n)$ and the probabilities for a family to have a boy or a girl are p and q respectively $(p + q = 1)$. A family is known to include exactly one boy. Find the probability that the boy is an only child. Assume that p is constant for all families.

15*. Box *A* contains 4 red balls and 8 blue balls and box *B* contains 6 red and 5 blue balls. Two ball are transferred from *A* to *B* and then *a* ball is withdrawn from *B*. (i) Find the chance that this ball is red. (ii) Given that the ball withdrawn is red, find the probability that *at least one* red ball was transferred to *B*.

16*. Urn A contains 9 cards numbered 1 through 9 and urn B contains five cards numbered 1 through 5. An urn is chosen at random and a card drawn. If the card shows an even number, another card is drawn from the same urn but if the card shows an odd number, a card is drawn from the other urn. What is the chance that both cards show (i) even numbers (ii) odd numbers ? If both cards show even numbers, what is the chance that they come from urn A.

17*. An urn contains 5 white and 5 black balls, four balls are drawn from this urn and put into another urn. From this second urn a ball is drawn and is found to be white. What is the chance of drawing a white ball at the next draw, the first white ball drawn is not replaced.

Exercise 2(d)

0. Show that if all of the probabilities are conditional on another event C, Bayes Theorem still remains true.

1. (a) A population consists of 60% males and 40% females, there being 5% men and 2% women are colour blind. A person chosen at random is found to be colour blind. What is the chance of this person being a male ?

(b) Five men out of 100 and 25 women out of 10,000 are colour blind. A colour blind person is chosen at random. Show that the probability of his being male is 0.8. Assume that male and female are in equal numbers.

2. (a) There are two identical boxes containing 4 white and 3 red balls; 3 white and 7 red balls. A box is chosen at random and a ball is drawn from it. Find the chance that the ball is white. If the ball is white prove that the probability that it is from first box is 40/61.

(b) Urn A contain 2 white, 1 black and 3 red balls. Urn B contains 3 white, 2 black and 4 red balls. Urn C contains 4 white, 3 black and 2 red balls. An urn is chosen at random and 2 balls are drawn. They happen to be red and black. Show that the chance that both balls came from urn B is 20/53.

3. (a) From an urn containing 3 white and 5 black balls, 4 balls are transferred into an empty urn. From this urn, 2 balls are taken and they both happen to be white. What is the probability that the third ball taken from the same urn will be white ? Consider the cases when the two balls drawn in the first draw are (a) returned, (b) not returned. **[Ans.** 7/12, 1/6]

(b) A box contains 5 chips, and of these it is equally likely that 0, 1, 2, 3, 4, 5 are white. A chip is drawn and it is found to be white. Show that P(this is the only white chip} = 1/15.

4. (a) Urns A_1, A_2 each contain 2 white and 3 black balls. Urns B_1B_2 each contain 1 white and 4 black balls. Urn C contains 4 white and 1 black ball. An urn is selected at random and a ball is drawn. It is found to be white. Find the probability that the ball comes out of the urns of the (i) second composition, (ii) third type. **[Ans.** 1/5, 2/5]

(b) There are 4 work-stationss of type A, each having 6 fitters and 3 turners and 3 work-stations of type B each having 2 fitters and 4 turners. One station is selected at random and a person is chosen at random from it. If he is a turner what is the probability that he came from type A station ? **[Ans.** 2/5]

5. Three machines A, B, C with capacities proportional to 2 : 3 : 4 [actual output 20,000, 30,000, 40,000] are producing bullets. The probabilities that these machines produce defectives are 0.1, 0.2, 0.1, respectively. One bullet is taken from day's production and found to be defective. Show that the probability that this bullet came from machine A or B is 5/6.

6. In a factory, two machines : M_1 and M_2 are used for the manufacture of screws, which may be uniquely classified as good or bad. M_1 produces per day n_1 boxes of screws of which on an average p_1% are bad, while the corresponding numbers of M_2 are n_2 and p_2. From the total production of both M_1 and M_2 for a certain day, a box is chosen at random, a screw taken out of it and it is found to be bad. Show that the chance that the selected box is manufactured by M_1 is $n_1p_1/(n_1p_1 + n_2p_2)$.
7. A building contractor receives bricks from three different suppliers : 35% from supplier A, 45% from supplier B and the rest from supplier C. 90% supply of A, 80% supply of B and 95% supply of C is according to specification. A brick drawn at random is not as per specification. Sho that the prob. that it came from B is 2/3.
8. From an urn containing 5 white and 5 black balls, 5 balls are transferred at random into an empty second urn from which one ball is now drawn. This is found to be white. Show that the probability that all 5 balls transferred from the first urn are white is 1/126.
9. There are five urns and they are numbered 1 to 5. Each urn contains 10 balls. Urn i has i defective balls and $10 - i$ non-defective balls, $i = 1, 2, ..., 5$. An urn is chosen at random and a ball is selected from it. What is the probability that a defective ball is selected ? If the selected ball is defective, show that the probability that it came from urn 5 is 1/3.
10. Bowl I contains 3 red chips and 7 blue chips, Bowl II contains 6 red chips and 4 blue chips. A bowl is selected at random and then one chip is drawn from this bowl. What is the chance that this chip is red ? If this chip is red find the conditional probability that it is drawn from Bowl II. If one chip is trnsferred from bowl I and placed in bowl II without seeing its colour and then one chip is withdrawn from bowl II, find the probability that it is *a* blue chip.
[**Ans.** 9/20, 2/3, 47/110]
11. Bowl A contains 6 red chips and 4 blue chips. Five of these 10 chips are selected at random and without replacement and put in bowl B which was originally empty. One chip is then drawn at random from bowl B. Relative to the hypothesis that this chip is blue, prove that the conditional probability that 2 red chips and 3 blue chips are transferred from bowl A to bowl B is 5/14.
12. A coin is tossed. If it turns up head, 2 balls will be drawn from urn A, otherwise 2 balls will be drawn from urn B. Urn A contains 3 black and 5 white balls. Urn B contains 7 black and 1 white ball. In both cases, selections are to be made with replacement. Prove that the chance that urn A is used, given that both the balls drawn are black, is $9(5)^6$ [**Ans.** $149 + 9(5)6$].
13. Mr. A chooses one of the coins C_1 (head-probability 3/4) and C_2 (head probability 1/4) and tosses the coin twice. Show that $P\{2 \text{ heads}\} = 5/16$, $P\{\text{heads only once}\} = 3/8$.
Instead of choosing C_1 or C_2, A chooses a fair coin and tosses it twice. State the procedure A should adopt so as to maximize the probability of at least one head.
14. An urn containing 5 balls has been filled up by taking 5 balls from another urn which originally had 5 white and 5 black balls. A ball is taken from the first urn and it is found to be white. Prove that the probability of drawing a white ball from among the remaining four is 4/9.
15. Bowl I contains 4 red and 5 black chips and bowl II contains 3 red and 6 black chips. One bowl is chosen at random a chip is drawn its colour noted and replaced back to the bowl. Again a chip is drawn from the same bowl, colour noted and restored to the bowl. This process is repeated four times, and the result yielded one red chip and 3 black chips. Prove that the probability that bowl I was chosen is 125/87.

16. (a) Three urns of the same appearance have the following composition : Urn I has 2 black and 1 white ball; urn II has 1 black and 2 white balls and urn III has 2 black and 2 white balls. One of the urns is selected at random and a ball is drawn. It turns out to be white. Prove that the chance of drawing a white ball again, if the first ball has not been returned, is 1/3.

(b) Urn A contains 1 white and 2 black balls, urn B contains 2 white and 1 black balls and urn C contains 3 black and 3 white balls. A cubical die is tossed. If 1, 2, 3 or 4 shows up, urn A is chosen, if 5 shows up urn B is chosen and if 6 shows up, urn C is chosen. A ball is drawn at random from the chosen urn. Obtain the probability that it is white.

17. Your host has arranged an unusual game for you and his other friends. Unseen by his friends his younger son spins a pointer which must stop in one of the three sectors, I of central angle of 144°, II of central angle 144° and III of central angle 72°. Also wunseen by his friends, his elder son throws a number of dice equal to the number of sectors in which the pointer stops. The sum of the spots is then announced, and the friends bet on the sector in which the pointer stopped. If the sum of the spots was 8, what are the odds for the various sectors.

18. Die A has 4 red and 2 white faces and Die B has 2 red and 4 white faces. A coin is flipped once. It if falls heads, the game continues by throwing Die A, if it falls tails, Die B is to be used. Show that the probability of getting a red face at any throw is 1/2. If the first two throws resulted in red, what is the probability of a red face at the third throw ? If the red turns up at the first n throws, show that the probability is $2^n(1 + 2^n)$ that Die A is being used.

19. From a packet containing 8 unused and 2 used blades, 5 are drawn at random and packed separately. From the second packet, a blade drawn randomly is found to be used. Find the prob. that it is the only used blade in the second packet. **[Ans.** 1/5]

20. Suppose a "five warning light" comes-on during flight in an aircraft. From past experience, the pilot has the following information :

(a) Probability that a fire will cause the warning light to come on = 0.95

(b) Probability that warning light will come on when there is no fire = 0.01.

(c) Probability that there will be a fire on any given mission = 0.001.

Show that the chance that there is actually a fire when the warning light is on, is 95/1094.

21. (a) Suppose that the reliability of chest X-ray test for the detection of T.B. is specified as follows : Of people with T.B., 90% of the X-ray tests detect the disease but 10% go undetected. Of people free of T.B. 99% of the X-rays are judged free of the disease, but 1% are diagnosed as showing T.B. from a large population of which only 0.1% have T.B. one person is selected at random, given a chest X-ray, and the radiologist reports the presence of T.B. What is the probability that the person has T.B. ?

(b) Mr. X is suspected to have one of the diseases D_1, D_2, D_3 which are rampant in society in the ratio 2 : 1 : 1. He is given a test which turns out to be positive in 25% of cases of D_1, 50% of D_2 and 90% of D_3. He is given these three tests and is declared positive in two. (event T). Show that $P(D_1 \mid T) = 18/69$, $P(D_2 \mid T) = 24/69$ and $P(D_3 \mid T) = 27/69$.

22. (a) A box contains $n_1 + n_2 + n_3$ tags where n_1 tags are numbered 1, n_2 tags are numbered 2 and n_3 tags are numbered 3. There are three urns, and they are numbered 1, 2, 3. Urn number j contains r_j red balls and b_j black balls. A tag is selected at random from the box, and then a ball is selected at random from the urn of the same number as the tag selected. If the ball selected is red, show that the chance that it came from urn number 2 is

$$n_2 r_2 [(r_2 + b^2)\, \Sigma_j\, r_j n_j / (r_j + b_j)]^{-1}.$$

(b) There are $n_1 + n_2 + ... + n_k = N$ urns, the first n_1 urns containing a_1 white, b_1 black balls each, the next n_2 urns contain a_2 white b_2 black balls each, ..., and the last n_k urns contains a_k

white and b_k black balls each. One urn is selected at random and a ball is drawn from it. If it is white find the prob. that the next ball drawn is also white, when (i) the first ball was replaced, (ii) the first ball is not replaced.

23. There is a series of n urns. In the kth urn, there are k white and $n-k$ black balls, $1 \le k \le n$. One urn is chosen at random and 2 balls are drawn from it. Both turn out to be white. Show that the prob. that the ith urn was chosen, where $3 < i < n$ is given by $i(i-1)\,[\Sigma k(k-1)]^{-1}$, $1 \le j \le n$.

24. Each of the n bags contains 4 red and 7 black chips while another bag contains 7 red and 4 black chips. A bag is chosen at random from the $(n+1)$ bags and 2 chips are drawn out of it together and both are found to be black. If it is now calculated that the probability that there are 7 red and 2 black chips remaining in the chosen bag is 1/15, show that $n = 4$.

25. A commuter residing in Dadar and working in Colaba must either return by Central Railway (C.R.) or Western Railway (W.R.) to get home. He varies his route, chosing the C.R. with probability 1/3 and W.R. with probability 2/3. If he goes by C.R. he gets home by 6 P.M. 75% of time. If he goes by W.R. he gets home by 6 P.M. only 70% of the time, but he likes the scenery better that way. If he gets home after 6 P.M. shows that the probability, that he used W.R. is 28/43.

26. The probability P_k that a family has k children is given by $P_k = (1-2p)/2^{k+1}$, $k \ge 2$. $P_0 = P_1 = p$. It is known that the family has 2 boys. Show that P {Family has only 2 children} = 27/64, P {Family has 2 girls as well} = 81/512.

27. A man is equally likely to choose any one of the three routes A, B, C from his house to the Railway station and his choise is not influenced by the weather. If the weather is fair, the chances of missing the train by routes A, B, C are respectively 1/20, 2/20, 4/20. He sets out on a fair day and misses the train. Show that the probability that the route chosen was C is 4/7. On a rainy day, the respective probabilities of missing the train by routes A, B, C are 1/20, 1/5, 1/2. On the average, one day in four is rainy. If he misses the train, what is the probability that the day was rainy ?

28. Buses of four routes A, B, C, D call at a certain bus-stop and out of every 14 buses, it can be assumed that 5 are A, 4 are B, 3 are C, and 2 are D. Two types of buses, old and new, are in service in equal numbers taking all 4 routes together, but the number of new buses on the different routes are such that if a person boards a new bus without noticing its routes indication, it is equally likely to be on any of the four routes. A person boards an old bus without noticing its route indication. Calculate the probability (a) from first principles, (b) by using Baye's theorem, that the bus is on route A.

29. Three prisoners A, B, C are told that any two of them are to be released on parole. A warder friend of prisoner A knows who are to be released. Prisoner A thinks of asking the warder for the name of one prisoner other than himself, who are to be released. He argues to himself that before he asks, his chances of release are 2/3. After knowing that B (say) shall be released, his own chances will fall to 1/2, because either A & B or B & C are to be released. And so A decides not to reduce his chances by asking. Show that A is mistaken in his calculations.

30. (a) A printing machine can print n letters say $\alpha_1, \alpha_2, ..., \alpha_n$. It is operated by electrical impulses, each letter being produced by a different impulse. Assume that p is the constant probability of printing the correct letter and the impulses are independent. One of the n impulses chosen at random, was fed into the machine twice and both times the letter α_1 was printed. Compute the probability that the impulse chosen was meant to print α_1.

[**Ans.** $P = (n-1)p_2/[np^2 - 2p + 1]$]

(b) Each packet contains N blades. Let P_k denote the probability that a packet contains k defective blades $(0 \le k \le m)$. A sample of n blades from a packet reveals $r \le m$ defective blades. Find the probability that the selected packet actually contains $k (\ge r)$ defective blades.

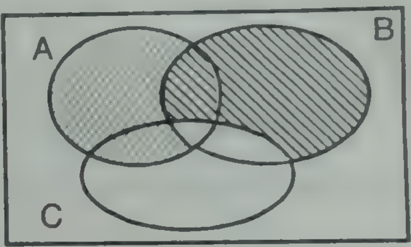
If a packet is rejected whenever it contains $d \le r$ defective blades, what is the chance that the packet is rejected ?

31. Assume that history goes back 2 million days. On each of these past days, the sum has risen. What is the probability that the sun will rise (i) tomorrow, (ii) each day for the next 2 million days.

32. A bag contains n balls, all possible proportions of white and black being equally likely. A ball is drawn from this bag and turns out to be white. This ball is repalced and then another ball is drawn, which also turns out to be white. If this ball is replaced, prove that the chance of getting a black ball in the next drawing is $\frac{1}{2}(n-1)(2n+1)^{-1}$.

33. In n trials, r successes have been observed. Assuming that aprori $P\{p = i/k\} = 1/k$, $1 \le i \le k$, find the probability that in the next n_1 trials, r_1 successes will occur on the assumption that aprori probability is uniformly distributed in (0, 1). Find also the limit as $n \to \infty$.

2-60. *Miscellaneous Worked-out Problems*



***Example* 1.** Show that $| P(A) - P(B) | \le P(A \Delta B)$.

Deduce : $P(A \Delta C) \le P(A \Delta B) + P(B \Delta C)$.

Solution. We estimate the difference as under :

$$\begin{aligned} P(A) - P(B) &\le P(A \cup B) - P(B) && [A \subseteq (A \cup B)] \\ &= P(A - B) && [P(C-D) = P(C) - P(D),\ D \subseteq C] \\ &\le P(A - B) + P(B - A) && [x \le x + y,\ x, y \in \mathbb{R}^+] \\ &= P(A \Delta B). \end{aligned}$$

Interchanging A and B gives $P(B) - P(A) \le P(A \Delta B)$

Combining the two components : $\quad | P(A) - P(B) | \le P(A \Delta B)$

Recall : $\quad A \Delta B = A\bar{B} \uplus \bar{A}B$.

Deduction. We use triangle inequality :

$$| P(A) - P(C) | \le | P(A) - P(B) | + | P(B) - P(C) |$$

$$\Rightarrow \quad P(A \Delta C) \le P(A \Delta B) + P(B \Delta C).$$

Note. Deduction trivially follows :

$$(A \Delta C) \subseteq (A \Delta B) \uplus (B \uplus C) \Rightarrow P(A \Delta C) \le P(A \Delta B) + P(B \Delta C).$$

***Example* 2.** An urn contains n white and m black balls, a second urn contains N white and M black balls. A ball is randomly transferred from the first to the second urn and then from the second urn to the first urn. If a ball is now selected randomly from the first urn, prove that the probability of its being white is,

$$\frac{n}{n+m} + \frac{mN - nM}{(n+m)^2 (N+M+1)}.$$

Solution. Denote a white ball by 'w', a black ball by 'b' and let wb represent a white ball taken from first urn and put into the second urn and a black ball returned to the first urn, etc. We can now distinguish four mutually exclusive and collectively exhaustive events according to the colour of the balls transferred as wb, bw, bb, ww. So,

$$P(wb) + P(bw) + P(bb) + P(ww) = 1 \quad ...(1)$$

Let W be the event of drawing a white ball from urn 1 after exchanges ; then by Multi-Stage rule :

$$P\{W\} = P(W \mid bw)\, P(bw) + P(W \mid wb)\, P(wb) + P(W \mid bb)\, P(bb) + P(W \mid ww)\, P(ww) \quad ...(2)$$

Now $\quad P(W \mid bw) = \dfrac{n+1}{m+n}$, $P(W \mid bb) = \dfrac{n}{n+m}$, $P(W \mid wb) = \dfrac{n-1}{m+n}$, $P(W \mid ww) = \dfrac{n}{m+n}$.

Substituting these values in (2), we get

$$(m + n)\, P(W) = (n + 1)\, P(bw) + (n - 1)\, P(wb) + nP(bb) + nP(ww)$$

$$= n[P(bw) + P(wb) + P(bb) + P(ww)] + P(bw) - P(wb) = n + P(bw) - P(wb) \quad [\text{by (1)}] \quad ...(3)$$

Now, $\quad P(bw) = P(b)\, P(w \mid b) = \dfrac{m}{m+n} \cdot \dfrac{N}{M+N+1}$ [black from I to II and white from II to I]

$P(wb) = P(w)\, P(b \mid w) = \dfrac{n}{m+n} \cdot \dfrac{M}{M+N+1}$ [white from I to II and black from II to I]

Substituting these values in (3), we get

$$P(W) = \frac{n}{m+n} + \frac{mN - nM}{(m+n)^2\,(M+N+1)}.$$

***Example* 3.** A marks a slip of paper with a plus sign with probability 1/3 and marks a minus sign with probability 2/3. A passes the slip to B who may or may not alter the sign before passing it to C. Next C passes the slip to D after perhaps chaing the sign. Finally D passes the slip to a referee after perhaps changing the sign. The referee finds a plus sign on the slip. It is known that B, C, D each changing sign with probability p. Compute $P\{A$ originally wrote a plus sign$\}$.

Solution. Let, $\quad T = \{A$ wrote a plus sign$\}$, $M = \{A$ wrote a minus sign$\}$

$\qquad R = \{$Referee found the plus sign on the slip$\}$

We shall use Multi-stage p-Rule :

$$P\{R\} = P\{T\}\, P(R \mid T) + P(M)\, P(R \mid M) \quad ...(1)$$

$P(R \mid T) = P\{$Plus sign was not changed $\uplus$ Changed exactly twice$\}$

$= P\{$plus sign not changed in 3 phases$\}$ + $P\{$sign changed twice in 3 phases$\}$

$$= \binom{3}{0} q^3 p^0 + \binom{3}{2} q p^2 = q(q^2 + 3p^2).$$

$P(R \mid M) = P\{$sign changed once $\uplus$ sign changed thrice$\}$

$= P($sign changed once$)$ + $P($sign changed thrice$)$

$$= \binom{3}{1} q^3 p + \binom{3}{3} q^0 p^3 = p(3q^2 + p^2).$$

$P(T) = \frac{1}{3}$, $P(M) = \frac{2}{3}$. Substituting into (i) gives

$$P\{R\} = \tfrac{1}{3}q(q^2+3p^2)+\tfrac{2}{3}p(3q^2+p^2)=\tfrac{1}{3}(q^3+2p^3+6pq^2+3qp^2)$$

$$= \tfrac{1}{3}[1+p^3+3pq^2],\ [(p+q)^3=1]$$

$\therefore\ P\{T \mid R\} = P\{R \mid T\}\, P(T)/P(R)$ (Reversal Identity)

$= q(q^2 + 3p^2)/(1 + p^3 + 3pq^2)$.

Note. When $p = 2/3$, $P\{T \mid R\} = 13/41$.

Problems with Solutions Provided at the End of the Text

1*. Two fair dice are thrown, the resulting outcomes being
$\Omega = \{(a, b) : 1 \le a \le 6,\ 1 \le b \le 6\}$. Three events A, B, C are defined as :
$A = \{(a, b) : a$ is odd$\}$, i.e. odd face with first die.
$B = \{(a, b) : b$ is odd$\}$, i.e. odd face with second die.
$C = \{(a, b) : a + b$ is odd$\}$, i.e. sum of points on two dice is odd.
Check whether A, B, C are independent or independent in pairs only.

2*. An urn contains two coins, one is honest and the other is 2-headed. A coin is selected at random and tossed. You are allowed to see the up-face which is heads. What is the chance that the hidden face is also heads ?

3*. A sportsman's chance of shooting an animal at a distance $r(> a)$ is a^2/r^2. He fires when $r = 2a$, and if he misses, he reloads and fires when $r = 3a, 4a, \ldots$ If he misses at distance na, the animal escapes. Find the odds against the sportman.

4*. Equally competent players A, B, C play a game. At each round the player who wins scores one point. The game is own by the *first* person who scores a total of 3 points. A wins the first round. Find the chance of B's winning.

Miscellaneous Exercises

1. (a) If $A \cap B = \varnothing$, show that, $\min\{P(A), P(B)\} = 0$.

(b) If A and B are exhaustive and independent events, such that $P(A) < 1$, show that $P(B) = 1$.

(c) If A, B, C are exhaustive and mutually independent events such that $P(A) < 1$, $P(B) < 1$, then $P(C) = 1$. Extend the result to four events.

(d) Show that $d(A, B) = P(A \,\Delta\, B)$ satisfies the triangle inequality and $| P(A) - P(B) | \le d(A, B)$.

2. (a) A proper event A is such that $P(A) = 1$. Prove that A is independent of every other event.

(b) Show that if an event A is independent of two mutually exclusive events B and C, then it is also independent of the event $B \cup C$.

(c) If A, B, C are mutually independent events, then $A \cup B$ and C are also independent.

(d) Show that if A, B, C are pairwise independent events and A is independent of $B \cup C$, then A, B, C are mutually indep.

3. (a) Show that A and B are independent iff $P(A \mid B) = P(A \mid \overline{B})$.

(b) If A, B, C are three events such that $P(C) = 0$, and A and B are independent, show that A, B, C are independent.

(c) Events B and C are independent, $P(B), P(C) > 0$, then for any event A,

$$P(A \mid B) = P(A \mid BC)\,P(C) + P(A \mid B\overline{C})\,P(\overline{C}).$$

Conversely, if this relation holds, $P(A \mid BC) \neq P(A \mid B)$, and $P(A) > 0$, then B and C are independent.

4. Consider the experiment of tossing two pair tetrahdra (with four faces numbered 1 through 4) and noting the numbers on the downturned faces.

(a) Give three proper events which are independent.

(b) Give three proper events which are only pairwise independent but not independent.

(c) Give four proper events which are independent. [Event A is proper if $0 < P(A) < 1$].

5. (a) Two dice (D_1 and D_2) are thrown. Find the probability of throwing (i) double, (ii) sum 5 or 9. If A is the set of events for which sum on the two dice is 11 and B is the set of events for which D_1 does not show 5, find $P(A \cap B)$ and show that A and B are not independent.

[**Ans.** $6\lambda, 8\lambda, \lambda, \lambda = 1/36$]

(b) A bag contains 5 white, 3 black and 4 red balls. A ball is drawn at random from the bag. Define the events B, W, R as follows :

B : The ball drawn is black, W : The ball drawn is white, R : The ball drawn is red. Examine pairwise and mutual independence of events B, W, R. Are B, W, R mutually exclusive ? Find $P(\overline{R})$, $P(B \cup W)$, $P(\overline{B} \cup \overline{W})$. [**Ans.** 8/12, 8/12, 1]

6. Urn A contains two white balls and two black balls ; urn B cntains three white and two black balls. One ball is transferred from A to B ; one ball is then drawn from B and it turns out to be white. Prove that the chance that the transferred ball was white is 4/7.

7. A white ball is dropped into a box containing n balls. What is the probability of drawing the white ball from this box if all the hypotheses about the initial colour composition of the balls are equally likely. [**Ans.** $p = (n+2)/2(n+1)$]

8. The probability that k calls are received at a telephone station during an interval of time t is equal to $P_t(k)$. Assuming that the number of calls during two adjacent intervals are independent, find the probability $P_{2t}(s)$ that s calls will be received during an interval $2t$.

[**Ans.** $p = \Sigma P_t(k)\, P_t(s-k), 0 \le k \le s$]

9. Each of k urns contains n identical balls, numbered 1 to n. One ball is drawn from each urn. Prove that $P\{$the greatest number drawn is $m\} = \{m^k - (m-1)^k]/n^k$ where $1 \le m \le n$.

10. The probability that a family has exactly k children is P_k ($\Sigma P_k = 1$). For any family size, all sex distributions are equilikely. What is the probability that a family has only one child, if it is given that the family has no girls (boys) ? [**Ans.** $1/2p_1 [\Sigma(1/2)^k p_k]^{-1}, 1 \le k < \infty$]

11. Let you classify people by 3 economic classes : Poor, Middle, Rich. Suppose you find 10% of the poor, 54% of the middle and 17% of the rich have colour television sets. Suppose further that the poor constitute 20%, the middle 75%, and the rich 5% of the population. A person selected at random is found to possess a colour T.V. set. How likely is he to be rich ? [**Ans.** $p = 0.0197$]

12. Among n persons $m(\le n)$ prizes are distributed by random drawing in turn from a box containing n tickets. Are the chances of winning equal for all participants ? When is it best to draw a ticket ? [**Ans.** $p = m/n$; immaterial]

13. Three ships A, B, C sail from England to India. Odds in favour of their arriving safely are 2 : 5, 3 : 7, and 6 : 11 respectively. Show that the chance that they all arrive safely is 18/595.

14. A, B, C each fire one shot at a target. The probabilities of their hitting the target are respectively given by 3/10, 1/4 and 1/10. If one bullet is found in the target, prove that the probability that it came from A's gun is 27/55.

15. Two cards are drawn in succession from a deck of playing cards. Determine (i) uncoditional prob. that the second card is an ace (the first card drawn being unknown) and (ii) the conditional prob. that the second card is an ace, if the first draw was an ace. [**Ans.** 1/15, 1/17]

16. A and B are two independent witnesses (i.e. there is no collusion between them) in a case. The probability that A will speak the truth is x and the probability that B will speak the truth is y. A and B agree in a certain statement. Show that the probability that this statement is true, is $xy/(1 - y - x + 2xy)$.

17. An urn contains a white and b black balls and a series of drawings of one ball at a time is made, the ball removed being returned to the urn immediately after the next draw is made. If P_n = probability that nth ball drawn is black, show that $P_n = (b - P_n - 1)/(a + b - 1)$. Hence find P_n.

18. The chance of success in each trial is p. If P_k is the probability that there are even number of successes in k trials, prove that $P_k = p + (1 - 2p)P_{k-1}$.

Deduce : $P_k = \frac{1}{2}[1 + (1 - 2p)^k]$.

19. If a day is dry, the conditional probability that the following day will also be dry is p. If a day is wet, the conditional probability that the following day will be dry is p'. If P_n is the probability that the nth day will be dry, prove that $P_n - (p - p')P_{n-1} - p' = 0,\ n \geq 2$. If the first day is dry, and $p = \frac{3}{4}$, $p' = \frac{1}{4}$, find P_n.

20. (a) A single die is tossed, then n coins are tossed where n is the number shown on the die. Find the probability of exactly 2 heads.

(b) A box contains 10 balls of which 5 are black. An integer n is chosen at random from the set $S = \{1, 2, 3, 4, 5, 6\}$ and then a sample of size n is drawn without replacement from the box. Find the chance that all the balls in the sample are black.

21. (a) A rook is palced on a specified square on a chess-board. At each move, it has an equal probability of moving to any one of the squares, other than the one currently occupied, in the same rank or file. Find the probability that it will return to (a) the original rank, (b) the original square at move number n.

(b) There are two lots of items. It is known that 90% of the items of one lot satisfy technical standards and 20% of the items of the other lot are defective. Suppose that an item from a lot selected at random turns out to be good. Show that the probability that a second item of the same lot will be defective, if the first item is returned to the lot after it has been checked is 5/34.

22. (a) Four barrels designated A, B, C, D contain balls of three different colours : white, black and red in proportion shown below :

Barrel	:	A	B	C	D
White	:	0.27	0.45	0.50	0.55
Black	:	0.13	0.35	0.35	0.25
Red	:	0.60	0.20	0.15	0.20

A barrel is chosen by spinning a six faced die with letters A, B, C, D and a ball is drawn from it. Show that the prob. that it is white or black is 0.75.

(b) An amplifier may burn one or both of its fuses if one or both of its tubes are defectives. Define the events : T_i = only tube i is defective, $i = 1, 2$; T_3 = both tubes are defective F_j = fuse j is burns out ($j = 1, 2$), F_3 = both fuses burn out. The various probabilities are as under : $P(F_1 \mid T_1) = 0.7$, $P(F_1 \mid T_2) = 0.3$, $P(F_1 \mid T_3) = 0.2$, $P(F_2 \mid T_1) = 0.2$, $P(F_2 \mid T_2) = 0.6$, $P(F_2 \mid T_3) = 0.02$, $P(F_3 \mid T_1) = 0.1$, $P(F_3 \mid T_2) = 0.1$, $P(F_3 \mid T_3) = 0.6$, $P(T_1) = 0.3$, $P(T_2) = 0.2$, $P(T_3) = 0.1$.
Evaluate $P(T_3 \mid F_3)$, $P(T_2 \mid F_3)$, $P(T_2 \mid F_1 \cup F_2)$.

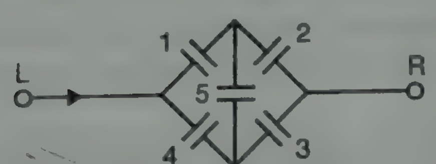
23. In a three game hockey play-off between teams A and B, team A has 58% chance of winning the first game. Winning a game boosts morale of either team so that its chances of winning its next game are impoved by 5%. Show that the probability that the team A wins the series is 0.61898. (Disregard ties).

24. (a) **Problem of four liars.** One person 'A' out of four receives information in the form of a 'yes' or 'no' and he transmits it to the second person 'B'. The second person transmits it to the third person 'C', the third to the fourth 'D' and the fourth communicates the received information to the sender. Given the fact that only one person in three tells the truth, show that the probability that the first liar told the truth if the fourth told the truth is 13/4]

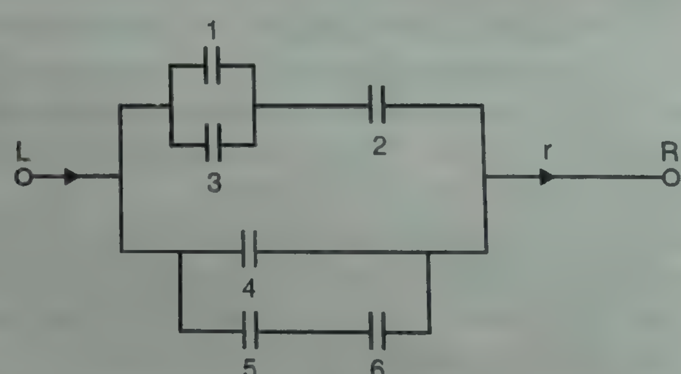
(b) It is known that each of 4 people A, B, C, D tells the truth in a given case with probability $\frac{1}{3}$. Suppose that A makes a statement, and then D says that C says that B says that A was telling the truth. Prove that the probability that A was actually telling the truth is 13/41.

25. In an election, candidate A received x votes and B receives y votes, with $x > y$. Assume the ballots are opened one by one, and all possible orders of votes (given the final results x and y) are equally likely. Show that, in the counting of votes.
$P\{A \text{ always ahead of } B\} = (x - y)/(x + y)$.

26. In the following figure (a) and (b) assume that the probability of a relay being closed is p and that a relay is open or closed is independently of any other. In each case find the probability that current flows from L to R.



(a)



(b)

[**Ans.** $2p^2(1 + p) - 5p^4 + 2p^5$; $p + 3p^2 - 4p^3 - p^4 + 3p^5 - p^6$]

If a man defrauds you one time, he is a rascal;
if he does it twice, you are a fool.

There is always room at the top for improvement. (Webster)

Appendix : Geometrical Methods for Probability Numericals

A

The application of geometry to problems of probability introduces in a natural way a broader concept of probability which goes beyond the elementary notion of the number of happenings of discrete events. Here, all points of interest lie within some prescribed region R. The probability measure is so defined that the total probability of the admissible region is unity, whereas that of all non-admissible region is zero. Since an uncountably infinite number of points lie in a finite region R of space S, the probability measure has to be defined in terms of the geometrical measure appropriate to the dimensionality (e.g. length, area, volume, etc). Thus,

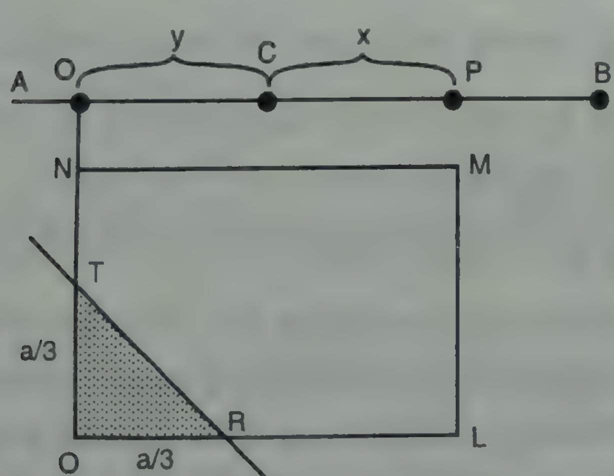
$$P\{x \in R\} = \text{Measure of region } R)/(\text{Measure of region } S) \; R \subseteq S.$$

It is generally assumed that the probability measure of R is directly proportional to geometrical measure of R. Geometrical problems often require the use of integration. The following problems illustrate the principles involved.

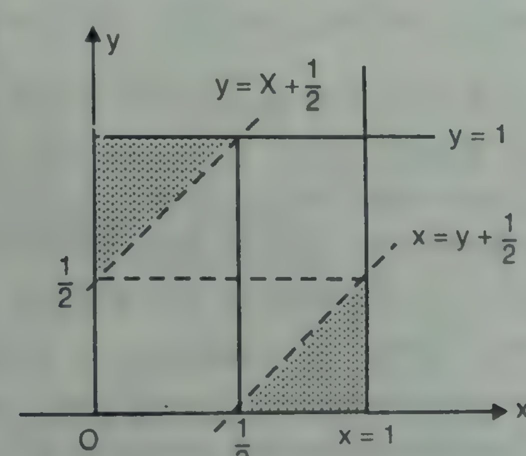
***Example* A-1.** Two points are selected in a line AB of length a so as to lie on opposite sides of its middle point C. Find the probability that the distance between them is less than $a/3$.

Solution. Choose points P and Q on the segments CB and CA and let $CP = x$, $CP = y$, so that the random choice provides $0 \leq x \leq a/2$, $0 \leq y \leq a/2$. We are interested in the event $x + y \leq a/3$. Draw the square $OLMN$ of side length $a/2$ and its sub-region triangle ORT, with $OR = OT = a/3$. The required probability is

$$p = \tfrac{1}{2}\cdot\left(\tfrac{1}{3}a\right)\left(\tfrac{1}{3}a\right)/\left(\tfrac{1}{2}a\right)\left(\tfrac{1}{2}a\right) = \tfrac{2}{9}.$$



Ex. A-1



Ex. A-2

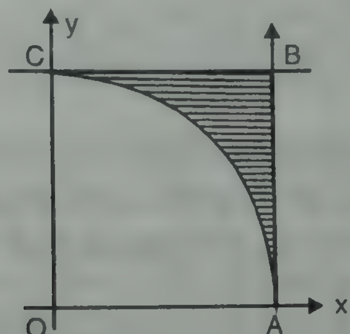
***Example* A-2.** Two points are chosen independently and at random from the interval [0, 1]. What is the chance that the two numbers differ by more than 1/2 ?

Solution. Let the numbers chosen be denoted by X and Y. Then their ranges are $0 \le x, y \le 1$. We need find $p = P\{\mid X - Y \mid \ge \frac{1}{2}\}$. This corresponds to the sum of two shaded areas of the unit square, obviously, $p = \frac{1}{4}$.

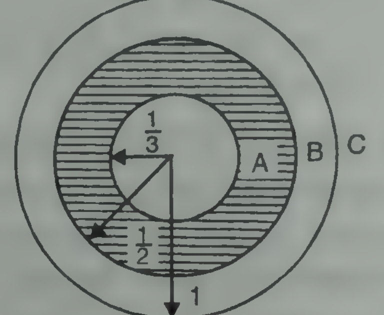
Example **A-3.** The outcomes of an experiment are represented by points in the square bounded by $x = 0$, $y = 0$, $x = 2$, $y = 2$ in the xy plane. If the probability is distributed unformly, determine the probability that $x^2 + y^2 > 4$.

Solution. We exhibit the sample space square of side 20 and the region $x^2 + y^2 > 4$, shown by light shade. The measure (area) of the shaded region is $4 - \pi$, hence the required probability measure $P\{x^2 + y^2 > 4\}$ is $p = [4 - \pi]/4 = 1 - (\pi/4)$. [See Figure]

Note. $P\{x^2 + y^2 > 1\} = 1 - P(x^2 + y^2 \le 1) = 1 - \frac{1}{4}\left(\frac{\pi}{4}\right) = 1 - \frac{\pi}{16}$.



Ex. A-3.



Ex. A-4.

Example **A-4.** A circular target is divided into three zones bounded by concentric circles of radii $\frac{1}{3}$, $\frac{1}{2}$, 1 respectively. If three shots are fired at random at the target, what is the probability that exactly one shot lands in each zone ?

Solution. The measure of sample space [area of circle of unit radius] is $\pi(1)^2$ i.e. π. Hence, the probab. of a shot falling in the region of radius $a (a < 1)$ is $p = \pi a^2/\pi = a^2$. Thus, the probability of a shot falling in region $A (a = \frac{1}{3})$ is $p_1 = \frac{1}{9}$, probability of a shot falling in region B is $p_2 = (\frac{1}{2})^2 - (\frac{1}{3})^2 = \frac{5}{36}$ and the probability of a shot falling in region C is $p_3 = 1 - (\frac{1}{2})^2 = \frac{3}{4}$. The shots fall in the three zones independents of each other. Hence by Product Rule, the probability of one arrangement, say first shot in A, second shot in B and third shot in C) is $p_1 p_2 p_3 = 5/432$. Since there are $3! = 6$ possible disjoint arrangements, the required probability is $6 \times 5/432 = 5/72$.

Example **A-5.** The sum of two positive quantities is equal to $2n$. Find the chance that the product of the two quantities is not less than $\frac{3}{4}$ times their greatest product.

Solution. Let $x > 0$, $y > 0$ be the given quantities so that $x + y = 2n$. To find the greatest value of xy, we notice that $z = xy = x(2n - x) = 2nx - x^2$, $z' = -2n - 2x$, $z'' = -2 < 0$. Now $z' = 0 \Rightarrow x = n$ and hence $y = n$ as $x + Y = 2n$.

Thus, the greatest value of xy is n^2. Now, we want,

$$p = P\{xy \not< (3n^2/4)\} = P\{xy > 3n^2/4\} \qquad \text{(Put } y = 2n - x)$$

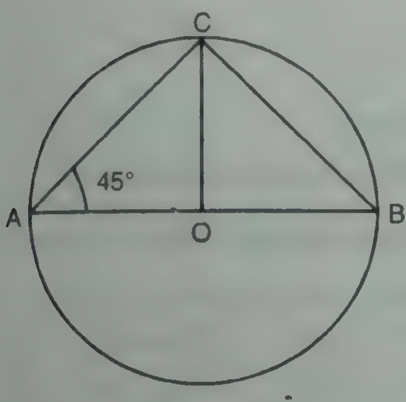
$$= P\{4x^2 - 8nx + 3n^2 \le 0\} = P\{(2x - 3n)(2x - n) \le 0\} = P[n/2 \le X \le 3n/2\}.$$

Since $0 \le x \le 2n$, (variation of x), total range of x values is $2n$.

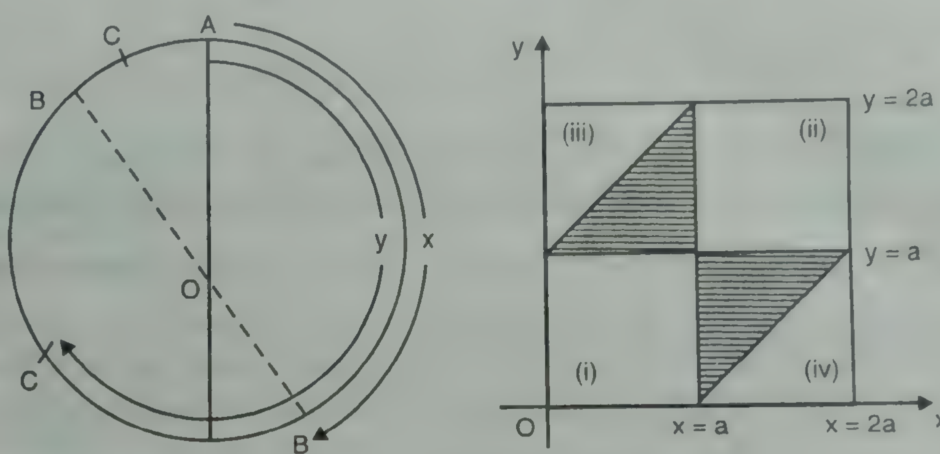
The favourable range is $(3n/2) - (n/2)$. It follows that, $p = (n/2n) = \frac{1}{2}$.

***Example* A-6.** AOB is a fixed diameter of a circle of radius a. A point is chosen at random within the circle. Find the chance that the point is within an inscribed isosceles triangle with AOB as one of the three sides.

Solution. If ABC is an isoscles triange $(AB = 2a, AC = BC = a\sqrt{2})$ its area is $\Delta = a^2$. There are two such triangles (two opposite sides of AOB), hence $p = 2a^2/\pi a^2 = 2/\pi$.



Ex. A-6.



Ex. A-7.

***Example* A-7.** Three points are taken at random on the circumference of a circle. Find the chance that they lie on the same semi-circle.

Solution. Let the length of the circumference of the given circle be $2a$. Fix a point A, let $AB = x$, $AC = y$, where we measure arc in the clockwise sense. The sampel space Ω is obviously.

$$\Omega = \{x : 0 < x < 2a,\ y : 0 < y < 2a\}.$$

Let $E \subset \Omega$ denote the set consisting of those points for which the points A, B, C lie on the same semi-circle. Then

$$E = \{x < a, y < a\} \uplus \{x > a, y > a\} \uplus \{x < a, y - x > a\} \uplus \{y < a, x - y > a\}$$

Notice that $x < a$ then $AB' = x + a$, C' exceeds AB' so that $y > x + a$.

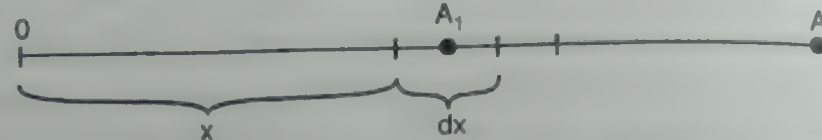
The measure favourable to E is shown on the indicated diagram.

$$\therefore \qquad p = \frac{\text{Measure of } E}{\text{Measure of } \Omega} = \frac{3a^2}{4a^2} = \frac{3}{4}.$$

***Example* A-8.** If $(n + 1)$ particles $A_1, A_2, \ldots, A_{n+1}$ be thrown at random upon a st. line OA (length a), each has the same chance of finding itself the $(k + 1)$th position in order reckoned from O to A. Also, since someone of them must occupy the $(k + 1)$th position, that chance is $1/(1 + n)$. Examine this rigorously.

Solution. The composite chance that a point $A_1 \notin [x, x + dx]$ and k unspecified points (particles) fall in $[O, OA_1]$ and rest $(n - k)$ fall in $[OA_1, OA]$ is

$$\binom{n}{k}\left(\frac{x}{a}\right)^k\left(1-\frac{x}{a}\right)^{n-k}\frac{dx}{a}.$$



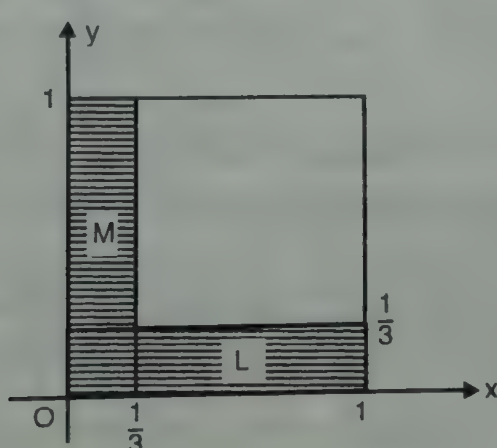
If B denotes the event that A_1 occupies the $(k + 1)$th place, irrespective of where it lies on OA, then

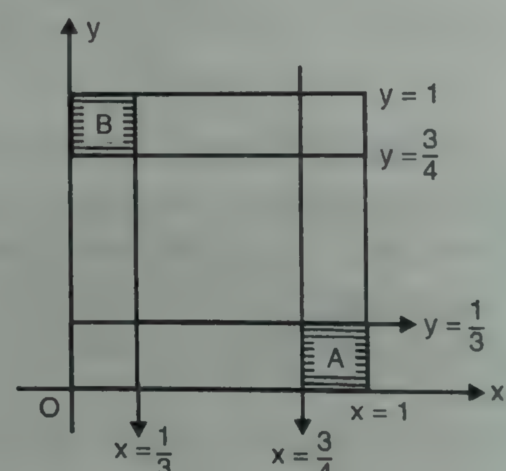
$$P(B) = \binom{n}{k}\int_0^a x^k\left(1-\frac{x}{a}\right)^{n-k}\frac{dx}{a^{k+1}} = 2\binom{n}{k}\int_0^{\pi/2}(\cos\theta)^{2(n-k)+1}(\sin\theta)^{2k+1}d\theta \left(\text{Put } \frac{x}{a} = \sin^2\theta\right) \quad ...(1)$$

$$= 2\frac{n!}{k!(n-k)!}\cdot\frac{\Gamma(n-k+1)\Gamma(k+1)}{2\Gamma(n+2)} = \frac{1}{n+1}.$$

Note. Put $x = at$ in (1), use beta integral for alternate evaluation.

***Example* A-9.** Two numbers are selected independently at random in the interval [0, 1]. If the smaller one is less than $\frac{1}{3}$, prove that the chance that the larger one is greater than $\frac{3}{4}$ is $p = \frac{3}{10}$.





Solution. Here $S = \{(x, y) : 0 \le x, y \le 1\}$; so that geometric configuration is a unit square. Let $M = \sup(x, y)$ and $m = \inf(x, y)$. We need find

$$p = P\{M \ge \tfrac{3}{4} \mid m \le \tfrac{1}{3}\} = P\{M \ge \tfrac{3}{4}, m \le \tfrac{1}{3}\}/P\{m \le \tfrac{1}{3}\}.$$

Now $\quad P\{m \le \frac{1}{3}\}$ = Shaded area $L \uplus M = \frac{1}{3} + \frac{1}{3} - \frac{1}{9} = \frac{5}{9}$.

$\quad P\{M \ge \frac{3}{4}, m \le \frac{1}{3}\}$ = Shaded areas A and B = $2 \times \frac{1}{3}.\frac{1}{4} = \frac{1}{6}$.

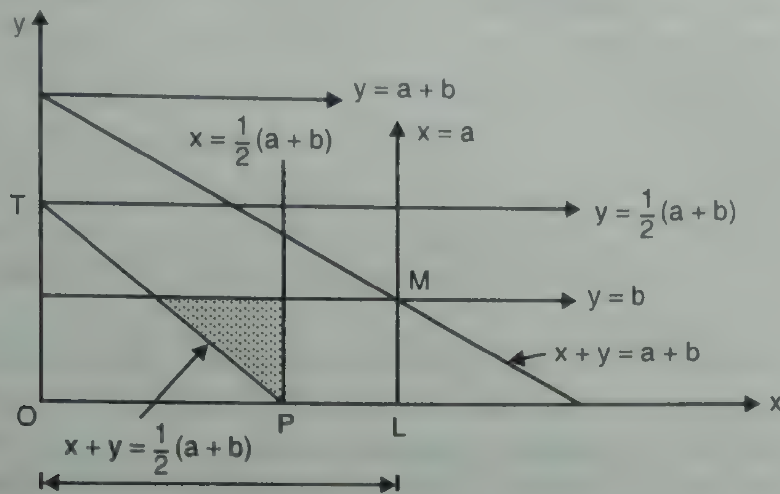
$\therefore\ p = (\frac{1}{6}/\frac{5}{9}) = \frac{3}{10}$.

***Example* A-10.** A rectilinear segment AB is divided by a point C into two parts : $AC = a$, $CB = b$. Two points X and Y are taken at random on AC and CB respectively. What is the probability that AX, XY, BY can form a triangle.

Solution. Let $AX = x$, $BY = y$, so that $XY = a + b - (x + y)$. The points x and y have uniform distribution in the intervals $[0, a]$ and $[0, b]$. Now a triangle is possible iff the sum of its two sides is greater than the third side. This fact supplies the conditions :

$$x + y > a + b - (x + y)\ ;\ x + [a + b - (x + y)] > y\ ;\ y + [a + b - (x + y)] > x$$

That is, $x + y > (a + b)/2$, $x < (a + b)/2$, $y < (a + b)/2$



We interpret geometrically these inequalities. The total region is the rectangle $OLMN$. The favourable region, governed by three inequalities, is the triangle PST, but only the triangular region PQR lies in the admissible region (sample space). Hence, the required probability p is

$$p = \text{measure of } \Delta PQR/\text{Measure of rect. } OLMN = \left(\tfrac{1}{2}b^2\right)/\,ab = b/2a.$$

Exercises for Geometrical Methods

1. Two points P and Q are taken at random on a given st. line OA of length a. Find the chance that the distance betwen them is (i) greater than a length $b < a$, (ii) less than a length $b(a)$. (iii) Find the probability that the point P is closer to Q than to O.

   [**Ans.** $p_1 = [1 - (b/a)]^2, p_2 = 1 - p_{12}, p_3 = \frac{1}{4}$]

2. Two points are selected at random on a given line of length a. Show that the chance that none of the three sections into which the line is thus divided is less than $a/4$ is 1/32.
3. The sides of a recntangle are taken at random each less than a and all lengths are equally likely. Show that the chance that the diagonal is less than a is $\pi/4$.
4. In each of two adjacent sides of a square of side a, a point is taken at random. Show that the chance that the length of the line joining the two points is between $a/2$ and a is $3\pi/4$.
5. A point is taken at random within an ellipse $(x^2/a^2) + (y^2/b^2) = 1$. Show that the chance that its distance from the centre exceeds b is $1 - (b/a)$.
6. A point is selected at random inside a circle. Prove that the chance that the point is closer to the centre of the circle than to its circumference is $\frac{1}{4}$.
7. Find the chance that a random triangle inscibed in a circle is (i) acute angled, (ii) obtuse angled, (iii) right-angled. [**Ans.** $p_1 = \frac{1}{4}, p_2 = \frac{3}{4}, p_3 = 0$]

8. A chord is drawn at random in a given circle. What is the probability that it is greater than the side of an equilateral triangle inscribed in that circle ? [**Ans.** $\frac{1}{3}$ and $\frac{1}{4}$, both possible]

9. A rod of length a is broken into three parts at random. Show that the probability that the three parts can form a triangle is $\frac{1}{4}$.

10. Bertrand Problem. *A* chord is drawn at random in *a* circle. Show that the chance that its length is not less than the radius of the circle, if it is assumed that all distances of the chord from the centre of the circle are equally likely, is $\sqrt{3}/2$.

11. A point is picked at random in a square. Show that the chance that it is lies inside the inscribed circle which is concentric with the square is $\pi/4$.

13. AB is a fixed diameter of a circle of radius a. *A* point is chosen at random within the circle. Show that the probability that the point is within an inscribed isosceles triangle with AB as one of the three sides is $2/\pi$.

13. A phonograph record 7" in diameter is thrown on the tiled floor of a game room. If the tiles are 10" square, shown that the chance that the record will lie wholy on one tile is 0.09.

14. A square sheet of tin, 20 centimetres wide, that contains 10 rows and 10 columns of circular holes, each 1 centimetre in diameter, with centres evenly spaced, at a distance of 2 centimetres apart. What is the chance that (i) a particle of sand (considered as a point) blown against the tin sheet will fall upon one of the holes and thus pass through it, (ii) a ball of diameter 0.5 cm thrown upon the sheet will pass through without hitting the tin sheet ? [**Ans.** $\pi/16$, $\pi/64$]

15. Show that the chance that the sum of two randomly chosen positive numbers, both ≤ 1, will not exceed 1, and their product will be $\leq 2/9$ is $(1/3) + (2/9) \ln 2$.

16. *Buffon's needle Problem.* A board is ruled with equidistant parallel lines, the width of strip between two consecutive parallel lines being d. A thin needle of length $l < d$ is thrown on the board. Show that the chance that the needle will intersect one of the lines is $l/\pi a$. What happens if $l > d$?

17. *Laplace Problem.* A board is covered with a set of congruent rectangales of dimensions a, b and a thin needle of length $l(l < a, l < b)$ is thrown on the board. Show that the chance that the needle will intersect a boundary of the rectangles is $[2(a + b)l - 12]/\pi ab$.

18. A floor is paved with tiles, each tile being a parallelogram such that the distance between pairs of opposite sides are a and b respectively, the length of the diagonal being l. A stick of length c falls on the floor parallel to the diagonal. Show that the probability that it will lie entirely on one tile is $[1 - (c/l)]^2$. If a circle of diameter d is thrown on the floor, show that the probability that it will lie on one tile is $[1 - (d/c)]\ [1 - (d/b)]$.

19. A bus of line A arrives at a station every 4 minutes and a bus of line B every 6 minutes. The length of an interval between the arrival of a bus of line A and a bus of line B may be any number of minutes from 0 to 4, all equally likely. Show that (a) $P\{$The first bus that arrives belongs to line $A\} = \frac{2}{3}$, (b) $P\{A$ bus of any line arrives within 2 minutes$\} = \frac{2}{3}$.

Teachers open the door. You enter by yourself. (Chinese Saying)

Genius may have its limitations, but stupidity is not thus handicapped.
(Elbert Hubbard)

Appendix : Sequences of Events B

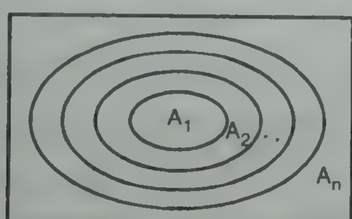
1. A sequence $(A_n : n \geq 1)$ is said to be an *Increasing* (Expanding) sequence of events if $A_1 \subset A_2 \subset \ldots \subset A_n \subset A_{n+1} \subset \ldots$

   Here we define a new event A_∞ by $A_\infty = \lim\limits_{n \to \infty} A_n = \bigcup\limits_{i=1}^{\infty} A_i$.

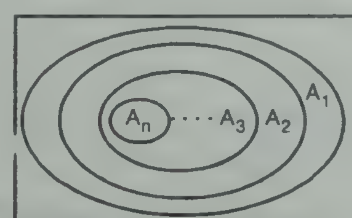
2. A sequence $(A_n : n \geq 1)$ is said to be a *Decreasing* (Contracting) sequence of events if

$$A_1 \supset A_2 \supset A_3 \supset \ldots \supset A_n \supset A_{n+1} \supset \ldots$$

   Here we define a new event A_∞ by $A_\infty = \lim\limits_{n \to \infty} A_n = \bigcap\limits_{i=1}^{\infty} A_i$.



Increasing sequence of events



Decreasing sequence of events

3. For any sequence of events. $\{A_n : n \geq 1\}$, we define the new events A^* and A^* as under :

$$A^* = \limsup_{n \to \infty} A_n = \bigcap_{i=1}^{\infty} \bigcup_{i=n}^{\infty} A_i = \{A_n, \text{i.o}\}, \text{ i.o.} = \text{infinitely of ten}$$

$$A_n = \liminf_{n \to \infty} A_n = \bigcup_{n=1}^{\infty} \bigcap_{i=n}^{\infty} A_i = \{A_n, \text{ult}\} \text{ ult } = \text{ ultimately}$$

A^* is called **limit superior** and A_* is called the **limit inferior** of the sequence (A_n) of events.

Note. $\{A_n, \text{f.o.}\} = \{A_n, \text{i.o.}\}^c$,   [f.o. = finitely often]

Remarks. A point $\omega \in A^*$ if, for all n, $\omega \in A \bigcup A_i$, $i = n, n + 1, \ldots \infty$. This means $\omega \in A_n$, $n = 1, 2, \ldots, \infty$. In other words, A^* consists of all points that are contained in an infinite number of the events A_i, $i \geq 1$.

A point $\omega \in A_*$ if for some n, $\omega \in \bigcap A_i$, $i = n, n + 1. \ldots \infty$. That is, if for some n, $\omega \in A_i$ for all i ¦ n. In other words, A_* consists of all points that are contained in all except a finitte number of the events A_i, $i \geq 1$. The event A_* occurs if all but a finite number of the A_i occur (the finite number of exceptions, indeed, varies with the elements of Ω, i.e. *its value is random*. The event A^* occurs iff infinitely many of the A_i occur. This reveals the strong requirement for the existence of $\lim A_n$.

4. If for any sequence of events $P\{A_n . n \geq 1\}$, $A^* = A_*$, then we say that $\lim A_n$ exists and define :

$$\lim_{n\to\infty} A_n = \lim_{n\to\infty} \sup A_n = \lim_{n\to\infty} \inf A_n.$$

5. For *a Real Sequence* we define the numbers :

$$\lim \sup (a_n) = \inf \{\sup\{a_n, a_{n+1}, \ldots\} : n = 1, 2, 3, \ldots\}$$
$$\lim \inf (a_n) = \sup \{\inf \{a_n, a_{n+1}, \ldots\} : n = 1, 2, 3, \ldots\}$$

Observe that : $\lim \inf (a_n) \leq \lim \sup (a_n)$.

Example (i) $(\lim \inf A_n)' = \lim \sup A'_n$. (ii) $\lim \sup (A_n \cup B_n) = (\lim \sup A_n) \cup (\lim \sup B_n)$.

B-1. *Continuity of Monotone Probability Function*

If $(A_n : n \geq 1)$ is either an increasing or decreasing sequence of events, whose limit is A, then

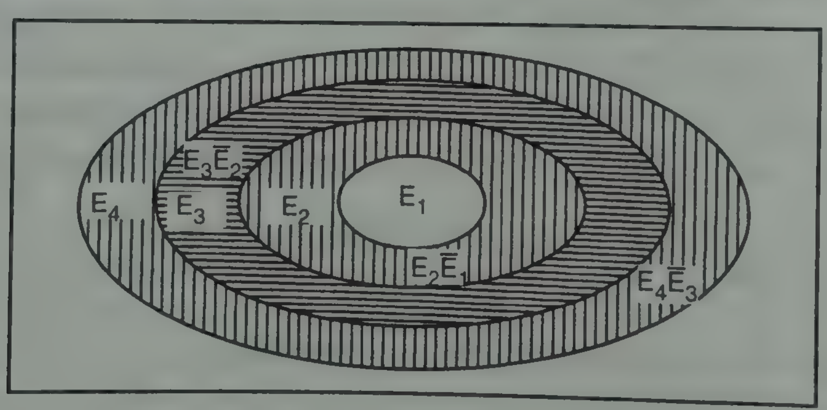
$$\lim_{n\to\infty} P(A_n) = P(\lim_{n\to\infty} A_n) = P(A).$$

Proof. By the monotonicity of probability, $\lim P(A_n)$ always exists. We now find this limit.

(i) Firstly, we consider the case of expanding (increasing) sequence $E_j \subset E_j + 1$, where by definition,

$$E = \lim_{n\to\infty} E = \bigcup_{j=1}^{\infty} E_j. \qquad \ldots(1)$$

We convert this non-disjoint union into *a* disjoint union as under. Let



$$B_1 = E_1, B_2 = (E_2 - E_1) = E_2\bar{E}_1, B_3 = E_3\bar{E}_2, \ldots, B_n = (E_n - E_{n-1}) = E_n\bar{E}_{n-1}.$$

Here each B_j is a part of E_j, not contained in any previous sets in the sequence. This B_j are disjoint sets and possess the obvious properties :

$$E_n = \bigcup_{i=1}^{n} E_j = \bigcup_{j=1}^{n} B_j, n \geq 1; \bigcup_{j=1}^{\infty} E_j = \lim_{n\to\infty} E_n = E_\infty \qquad \ldots(2)$$

Now, we use axioms of finite and infinite σ-additivity :

$$P(\bigcup_{i=1}^{\infty} E_i) = P(\bigcup_{j=1}^{\infty} B_j) = \sum_{j=1}^{\infty} P(B_j) = \lim_{n\to\infty}\left[\sum_{j=1}^{n} P(B_j)\right] = \lim_{n\to\infty} P(\bigcup_{j=1}^{n} B_j) = \lim_{n\to\infty} P(E_n).$$

Using (1), this amount to :

$$P(\lim_{n\to\infty} E_n) = \lim_{n\to\infty} P(E_n). \qquad ...(3)$$

(ii) Now if $\{E_n, n \geq 1\}$ is a decreasing sequence, then $\{\overline{E}_n, n \geq 1\}$ is an increasing sequence. Hence by (3), we have

$$P\left\{\bigcup_{i=1}^{\infty} \overline{E}_i\right\} = \lim_{n\to\infty} P(\overline{E}_n).$$

From $\bigcup_{i=1}^{\infty} \overline{E}_i = \left(\bigcap_{i=1}^{\infty} E_i\right)^c$ we get $P\left[\bigcap_{i=1}^{\infty} E_i\right]^c = \lim_{n\to\infty} P(\overline{E}_n)$

By nogation Rule, $1 - P\left\{\bigcap_{i=1}^{\infty} E_i\right\} = \lim_{n\to\infty} [1 - P(E_n)] \Rightarrow P\left(\bigcap_{i=1}^{\infty} E_i\right) = \lim_{n\to\infty} P(E_n)$

Note. We can also establish (ii) without using (i). See Example **B-3.**

B-2. *Continuity of General Probability Function*

If $(A_n : 1 \leq n \leq \infty)$ is any sequence of events, then

$$P(A^*) \leq \liminf_{n\to\infty} P(A_n) \leq \limsup_{n\to\infty} P(A_n) \leq P(A^*). \qquad ...(1)$$

If $\lim A_n$ exists, then $P(\lim_{n\to\infty} A_n) = \lim_{n\to\infty} P(A_n)$ ...(2)

Proof. Note that for increasing and decreasing sets of events, $E_n \subset E_{n+1}$, $D_{n+1} \subset D_n$. To simplify writing, we set

$$E_n = \bigcap_{i=n}^{\infty} A_i, D_n = \bigcup_{i=1}^{n} A_i. \qquad ...(i)$$

Thus, $\{E_n\}$ is an increasing sequence, and $\{D_n\}$ is a decreasing (contracting) sequence of events.

Hence, by continuity of monotone probability function

$$P(A_*) = \lim_{n\to\infty} P(E_n) = \sup_n P(E_n) \ ...(ii) \quad P(A^*) = \lim_{n\to\infty} P(D_n) = \sup_n P(D_n) \qquad ...(iii)$$

From (i), $A_j \supset E_n \ \forall\, j \geq n$; $A_j \subset D_n \ \forall\, j \geq n$, hence

$$P(E_n) \leq P(A_j) \Rightarrow P(E) \leq \inf P(A_j), j \leq n \qquad ...(iv)$$

$$P(D_n) \geq P(A_j) \Rightarrow P(D_n) \geq \sup P(A_j), j \geq n \qquad ...(v)$$

From (ii) and (iv) as well as (iii) and (v), we obtain,

$$P(A_*) \leq \sup_n \inf_{j\geq n} P(A_j) = \liminf_n P(A_n), \quad P(A^*) \geq \inf_n \sup_{j\geq n} P(A_j) = \limsup_n P(A_n) \qquad ...(vi)$$

Since, $\liminf (a_n) \leq \limsup (a_n)$, we get from (vi) the result (1).

If $\lim A_n$ exists $(A^* = A_*)$, then $P(\lim_{n\to\infty} A_n) = \lim_{n\to\infty} P(A_n)$.

This is obvious, because the limit of a monotone class of events always exists.

Remarks. Continuity of probability measure shows that finite additivity yields σ-addivity.

B-3. *First Borel-Cantelli Lemma*

Let $A_1, A_2, \ldots$ be events (not necessarily independent) in a given probability space and let $P(A_i) = P_i$. If ΣP_i, $(1 \le i < \infty)$ is convergent, then $P(A^*) = 0$.

Proof. Observe that : $A^* = \limsup_n A_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k$. For any n, since $A^* \subset \bigcup_{k=n}^{\infty} A_k$, we have

$$0 \le P(A^*) \le P\left(\bigcup_{k=n}^{\infty} A_k\right) \le \sum_{k=n}^{\infty} P(A_k) = \sum_{k=n}^{\infty} P_k \qquad \ldots(1)$$

where we have used Monotone law and Boole's inequality to get (1). Now, since ΣP_i is convergent to p (say), we have

$$p = \lim_{n\to\infty} \sum_{i=1}^{n} P_i \Rightarrow \lim_{n\to\infty}\left(p - \sum_{i=1}^{n} P_i\right) = 0 = \lim_{n\to\infty} \sum_{i=n+1}^{\infty} P_i .$$

Hence, the last member of Eq. (1) tends to zero as $n \to \infty$. It follows by squeeze rule that, $P(A^*) = 0$.

B-4. *Second Borel-Cantelli Lemma*

Let $A_1, A_2, \ldots$ be events which are mutually independent, and let $P(A_i) = P_i$. If ΣP_i, $(1 \le i < \infty)$ is unbounded (divergent), then $P(A^*) = 1$.

Proof. Observe that since $(A^*) = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k$, hence : $(A^*)^c = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} \overline{A}_k = \bigcup_{n=1}^{\infty} B_n$ (say),

with obvious notation. Now

$$0 \le P[(A^*)^c] = P\left(\bigcup_{n=1}^{\infty} B_n\right) \le \sum_{n=1}^{\infty} P(B_n), \qquad \text{[by Boole's inequality]} \ldots(2)$$

Since the events A_i are mutually independent, so are $\overline{A}_i$, and hence

$$P(B_n) = \prod_{i=n}^{\infty} P(\overline{A}_i) = \prod_{i=n}^{\infty} [1 - P(A_i)] = \prod_{i=n}^{\infty} Q_i . \qquad [Q_i = 1 - P_i]$$

To obtain a bound on this, we notice that $1 - x < e^{-x}$, $0 < x < 1 \Rightarrow Q_i < \exp(-P_i)$, $(x = P_i)$.

Thus : $\prod_{i=n}^{N} Q_i \le \exp\left(-\sum_{i=n}^{N} P_i\right)$ where $N > n$.

As $N \to \infty$, $\sum_{i=n}^{N} P_i \to \infty$, $\exp\left(-\sum_{i=n}^{N} P_i\right) \to 0 \Rightarrow \prod_{i=n}^{\infty} Q_i = 0$.

Thus, $P(B_n) = 0$, and Eq. (2) provides that $P[(A^*)^c] = 0$. Whence $P(A^*) = 1$ by complement rule.

Cor. Kolmogorov Zero-one Law

For independent sequence of events $\langle A_n \rangle$, if an event is, in its tail field, then

$P(A^*) = 0$, if $\Sigma P_i < \infty$ and $P(A^*) = 1$ if $\Sigma P_i = \infty$.

Thus, $P\{A_i$ occur infinitely often$\} = 0$, or 1. No other value is possible.

This explains why the term zero-one law is often used for Borel-Cantelli lemmas.

Note. When $\Sigma P_i < \infty$, then $P\{A_i$ occur finitely often$\} = 1$.

B-5. *Illustrations for Sequences of Events*

***Example* B-1.** State precisely a set of axioms of probability and deduce the classical rule for the calculaton of probabilities.

Solution. For aximons, see Art. 1-51, p. 21.

Let $\Omega = \{\omega_1, \omega_2, ...\}$ be a finite or countably infinite set. Let f be a real-valued function defined on Ω that satisfies

$$f(\omega) \geq 0, \forall\ \omega \in \Omega \text{ and } \Sigma_\Omega f(\omega) = 1. \qquad ...(1)$$

Then, we can define a function P on the class F of all subsets of Ω by letting

$$P(A) = \Sigma_A f(\omega), \forall\ A \subset \Omega \qquad ...(2)$$

We have $P(A) \geq 0$ and $P(A) \leq P(\Omega)$ for every A, since $f(\omega) \geq 0$, $\forall\ \omega$ and $P(\Omega) = 1$ [by (1)] Further if $A \cap B = \varnothing$, then

$$P(A \uplus B) = \sum_{A \cup B} f(s) = \sum_A f(\omega) + \sum_B f(\omega) = P(A) + P(B).$$

Thus, axioms P_3' (and similarly P_3) is satisfied. It follows that P is a probability function and $(\Omega, \mathcal{F}, P)$ is a probability space.

Taking $A = \{\omega\}$ as singleton, in (2) yields $P(\{\omega\}) = f(\omega)$ for $\omega \in \Omega$. Thus, $f(\omega)$ gives the probability that the experimental outcome is ω.

Note. If Ω is finite, and if $f(\omega) = 1/|\Omega|$ for all $\omega \in \Omega$, then (2) yields $P(A) = |A|/|\Omega|$ for $A \subset \Omega$, which is the classical formula for evaluating the probabilities.

***Example* B-2.** Show that the σ-additivity follows from the continuity of probabilty function, i.e. continuity of probability is equivalent to σ-additivity.

Solution. Let the probability function P defined on F be continuous. Consider a sequence of m.e. events $A_1, A_2, ...$ and define

$$A = \bigcup_{k=1}^{\infty} A_k - \left(\bigcup_{k=1}^{n} A_k\right) \cup \left(\bigcup_{k=n+1}^{\infty} A_k\right) = E_n \cup D_n;\ E_n = \bigcup_{k=1}^{n} A_k,\ D_n = \bigcup_{k=n+1}^{\infty} A_k \qquad ...(1)$$

Obviously, $\{E_n\}$ is an increasing sequence of events with $\lim E_n = A$ and $\{D_n\}$ is a decreasing sequence of events with $\lim D_n = \varnothing$ (since A_j are disjoint). Since P is assumed to be continuous,

$$\lim P(D_n) = P(\lim D_n) = P(\varnothing) = 0.\ (n \to \infty) \qquad ...(2)$$

Using finite additivity of P-measure, we get from (1)

$$P(A) = P(E_n) + P(D_n) = \sum_{k=1}^{n} P(A_k) + P(D_n)$$

As $n \to \infty$, the above, in virtue of (2), reduces to

$$P\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} P(A_k). \qquad \text{(σ-additivity axiom)}$$

***Example* B-3.** If $\{D_n : 1 \le n < \infty\}$ is a decreasing sequence of events whose limit is the event D, then prove that $\lim P(D_n) = P(D)$, as $n \to \infty$.

Solution. This theorem is already proved in Art. B-1. The following is an independent proof. When the D_n are non-increasing, we want to show that (Draw Figure) :

If $D = \bigcap_{n=1}^{\infty} D_n$, $\lim_{n\to\infty} (D_n) = P(D)$. Decompose D_n into the unions of disjoint events as follows :

$$D_n = \left(\bigcup_{j=n}^{\infty} D_j \bar{D}_{j+1}\right) \cup D. \qquad ...(2)$$

Observe that,

$DD_j\bar{D}_{j+1} = \varnothing$, $1 \le j < \infty$, because $D = \cap D_n (1 \le n < \infty)$; hence $D \cup D_j\bar{D}_{j+1} = \varnothing$ $(n \le j < \infty)$.

Now by σ-additivity of P, (2) gives

$$P(D_n) = \Sigma P(D_j \bar{D}_{j+1}) + P(D) \qquad (j = n \text{ to } \infty) \qquad ...(3)$$

From $\cup D_j\bar{D}_{j+1} \subseteq \Omega$ ($j = 1$ to ∞), we get $\Sigma P(D_j\bar{D}_{j+1}) \le 1$, so

$$\sum_{i=1}^{\infty} P(D_j\bar{D}_{j+1}) + \sum_{j=n}^{\infty} P(D_j \bar{D}_{j+1}) \le 1.$$

$$\therefore \quad \lim_{n\to\infty} \sum_{i=1}^{n} P(D_j \bar{D}_{j+1}) + \lim_{n\to\infty} \sum_{j=n}^{\infty} P(D_j\bar{D}_{j+1}) \le 1, \Rightarrow \lim_{n\to\infty} \sum_{j=n}^{\infty} P(D_j\bar{D}_{j+1}) = 0. \qquad ...(4)$$

Using (4) into (3) provides (1).

***Example* B-4.** Suppose $\{E_n, n \ge 1\}$ and $\{H_n, n \ge 1\}$ are increasing sequences of events having limits E and H. Prove that if E_n is indep. of H_n $\forall$ n, then E is indep. of H.

Solution. Let $G_n = E_n \cap H_n - E_n H_n$. Since E_n and H_n are monotone increasing, so G_n is also monotone increasing. Since probability function is continuous, hence

$$\lim P\{E_n H_n\} = P\{\lim E_n \lim H_n\} = P(E\,H).$$

***Example* B-5.** For sequences $\langle A_n \rangle$ and $\langle B_n \rangle$ of events on some sample space S, let $\lim_{n\to\infty} P(B_n) = 1$. Show that $\lim_{n\to\infty}\{P(A_n \mid B_n) - P(A_n)\} = 0$. Deduce that if $\lim_{n\to\infty} P(A_n \mid B_n) < \infty$ or $\lim_{n\to\infty} P(A_n) < \infty$, both limits exist and $\lim_{n\to\infty} P(A_n \mid B_n) = \lim_{n\to\infty} P(A_n)$.

Solution. Since $\lim_{n\to\infty} P(B_n) = 1$, $\lim [1 - P(B_n)] = 0$. Further, $P(A_n \cup B_n) \ge P(B_n)$ yields; $\lim P(A_n \cup B_n) = 1$.

$$\text{Now } P(A_n \mid B_n) - P(A_n) = \frac{P(A_n B_n)}{P(B_n)} - P(A_n) = \frac{P(A_n B_n) - P(A_n)P(B_n)}{P(B_n)} = \frac{P(A_n) + P(B_n) - P(A_n \cup B_n) - P(A_n)\,P(B_n)}{P(B_n)}$$

$$\therefore \quad \lim_{n\to\infty} \{P(A_n \mid B_n) - P(A_n)\} = \frac{\lim |(P(A_n)[1 - P(B_n)] + P(B_n) - P(A_n \cup B_n)|}{\lim P(B_n)} = \frac{0 + 1 - 1}{1} = 0. \quad ...(1)$$

From elementary calculus of real sequences, if $c_n = a_n - b_n$ possesses a limit, then so do $<a_n>$ and $<b_n>$ iff at least one of these does possess a limit. Consequently, if either $\lim P(A_n \mid B_n) < \infty$ or $\lim P(A_n) < \infty$, then both do possess finite limits and by (1).

$$\lim P(A_n \mid B_n) = \lim P(A_n), \; (n \to \infty).$$

Results without causes are much more impressive. (Sherlock Holmes)

[The Stock-Brokers Clerk]

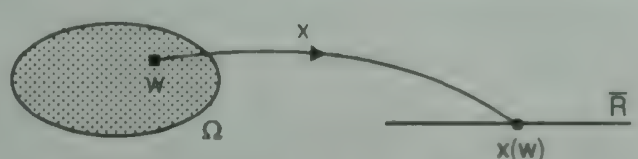
He talks at random. Sure the main is mad. (Shakespeare : Henry VI)

One-Variate Distribution Theory

3

3-10. *Random Variables*

In random experiments, we are interested in the *numerical outcomes* (i.e. numbers associated with the outcomes) of the experiment. For example, we ask for the sum of the points on the pips when a pair of dice is thrown. When 50 coins are tossed, we ask for the *number* of heads and so on. Whenever we associate a real number with each outcome of a trial, we are dealing with a *function* whose *range* is the set of real numbers we ask for. Such a function is called a *random variable* (r.v.) chance variable, stochastic variable or simply a **variate**. We shall consider only real-valued random variables in this book.



***Definition* 1.** Let S be a denumerable set (i.e. finite or countably infinite). A **discrete** random variable X on a probability space $(\Omega, \mathcal{F}, P)$ is a real-valued function X mapping domain Ω into range $S = \{x_1, x_2, ..., x_n, ...\}$ of the real numbers R such that $\{\omega : X(\omega) = x_i\} \in \mathcal{F}$, for all i.

The requirement : $[\omega : X(\omega) = x_i\} \in \mathcal{F}$ allows us to define $p(x) = P(X = x)$. The collection $\{p(x) : x \in S\}$ is called the *distribution* of X and it satisfies the obvious :

$$0 \le p(x) \le 1, \quad \sum_{x \in S} p(x) = 1.$$

Note. Discreteness also implies that $\mathcal{F}$ contains all subsets of Ω, in particular, subsets of the form $[\omega : X(\omega) \le x\}$.

***Definition* 2.** A random variable X on a probability space $(\Omega, \mathcal{F}, P)$ is a real-valued function X defined on Ω, such that for $-\infty < x < \infty$, $\{\omega : X(\omega) \le x\} \in \mathcal{F}$. Thus $X : \Omega \to R$ such that $\{\omega : X(\omega) \le x\} \in \mathcal{F}, \forall\, x \in R$.

Terminology. Henceforth, we shall always write $X = t$ for $X(\omega) = t$ and $\{X \le x\}$ for $\{\omega : X(\omega) \le x\}$.

Remark. The definition of a random variable does not require the notion of Probability Function. Thus, if X is a r.v. on $(\Omega, \mathcal{F}, P\}$ and the probability measure is changed, then X is still a r.v. It should also be clear from the above definition that every real-valued function need not be a random variable (vide Example 1). Whether a given real-valued function is a random variable or not, depends on the underlying σ-field.

Illustrations (a) Toss of a single coin : $\Omega = \{h, t\}$. Let X be the r.v. defined by : $X(h) = 1$, $X(t) = 0$; i.e. $X = 1$, if head occurs ; $X = 0$, if tail occurs.

(b) Roll of a single die : $\Omega = \{1, 2, 3, 4, 5, 6\}$. Let $X(\omega) = \omega$, $\omega = 1, 2, \ldots, 6$; i.e. the variate X denotes the value of the upturned face.

(c) Two distinct dice are rolled : $\Omega = \{(a, b) : a, b = 1, 2, ..., 6\}$. Let

$$X(a, b) = a + b,\ 1 \le a, b \le 6.$$

The r.v. X denotes the sum of two numbers upturned when two dice are thrown.

3-11. *Random Variable Versus Real-valued Function*

If Ω is a sample space of a random experiment and if $\mathcal{F}$ is the σ-field of all subsets of Ω, then every real-valued function X on Ω is a random variable.

Proof. Let $X : \Omega \to R$ be any real-valued function. For any real number x, the set $\{\omega : X(\omega) \le x\}$ is a subset of Ω. Since $\mathcal{F}$ contains all subsets of Ω, so $\{\omega : X(\omega) \le x\} \in \mathcal{F}$, for all $x \in R$. Hence X is a random variable.

Remarks. In the case of discrete sample spaces, the underlying σ-field will be taken as the set consisting of all subsets of sample space (i.e. Power Set). In the case of sample spaces which contain uncountably many points (e.g. R) the σ-field $\mathcal{F}$ is generally taken as the smallest σ-field containing all the semi-closed intervals of the form $]\,a, b\,]$. This smallest σ-field containing the intervals of the form $]\,a, b\,]$ is known as Borel σ-field.

Example 1. Give an example of a variable which is not a r.v.

Solution. Let $\Omega = \{a, b, c, d\}$ be a sample space with associated σ-field $\mathcal{F} = \{\varnothing, \{a\}, \{b, c, d\}, \Omega\}$. Define $X : \Omega \to R$ by $X(a) = 0 = X(b)$, $X(c) = X(d) = 2$.

X is a real-valued function on Ω. Now consider the set $\{\omega : X(\omega) \le 1\} = \{a, b\} \notin \mathcal{F}$.

By definition of a random variable, x is *not* a random variable.

Example 2. Let Ω (sample space) be the set of all real numbers i.e. $\Omega = R$. Suppose that $\mathcal{F}$ $\{A : A$ is countable or $\bar{A}$ is countable$\}$, is the σ-field associated with Ω.

(a) Let $X : \Omega \to R$ be defined by $X(\omega) = \omega$, for all $\omega \in \Omega$. Then X is not a random variable.

(b) Let $Y : \Omega \to R$ be defined by $Y(\omega) = 1$, if ω is irrational ; $Y(\omega) = 0$, if ω is rational. Then, Y is a random variable.

(c) Let $Z : \Omega \to R$ be defined by : $Z(\omega) = 0$, if ω is rational.

$Z(\omega) = -1$, if ω is a negative irrational ; $Z(\omega) = 0$, if ω is a positive irrational, then, Z is not a random variable.

Solution. (a) By definition, $X : \Omega \to R$ is such that $X(\omega) = \omega\ \forall\ \omega \in \Omega = R$. Consider the set $A = \{\omega : X(\omega) \le 1\} =]\,\infty, 1\,]$. Then $\bar{A} =]\,1, \infty\,[$. Since both A and $\bar{A}$ are uncountable, the function X is not a random variable

(b) Let Q denote the set of all rational numbers. Since Q is countable, $Q \in \mathcal{F}$. Let y be the real number and consider the set $A = \{\omega : Y(\omega) \le y\}$.

If $y < 0$, $A = \varnothing \in \mathcal{F}$; if $0 \le y < 1$, $A = Q \in \mathcal{F}$; and if $y \ge 1$, $A = \Omega \in \mathcal{F}$.

If follows that Y is a random variable.

(c) For the function Z, consider the set $A = \{\omega : Z(\omega) \leq -\frac{1}{3}\}$.

By definition of Z, A is the set of all negative irrational numbers. A is not countable and its complement $\overline{A}$ is also not countable, because $\overline{A}$ consists of rational and positive irrationals and is thus uncountable. So $A \notin \mathcal{F}$ and therefore, Z is not a random variable.

Remark. Since $|Z| = Y$ and Y is a random variable, so $|Z|$ is a random variable although Z is not a random variable.

Note. It is easy to check that $\mathcal{F}$ is a σ-field.

Constant R.V. Theorem

Let $\{\Omega, \mathcal{F}, P\}$ be an arbitrary probability space. Let $c \in R$ be any real number. Then, the constant function $X : \Omega \to R$, defined by $X(\omega) = c$, $\forall\ \omega \in \Omega$ is a random variable.

Proof. Let $X : \Omega \to R$ be defined by $X(\omega) = c$, $\forall\ \omega \in \Omega$. If $x \in R$ is any real number, then $A = \{\omega : X(\omega) \leq x\} = \{\omega : c < x\}$. If $x < c$, then $A = \varnothing \in \mathcal{F}$; if $x \geq c$, $A = \Omega \in \mathcal{F}$.

Thus, X is a random variable.

Degenerate Variate. Let $(\Omega, \mathcal{F}, P)$ be an arbitrary probability space. A constant function X defined by $X(\omega) = c$, $\forall\ \omega \in \Omega$, is called a *Degenerate random variable*.

3-12. *Indicator Random Variable and its Properties*

Let $(\Omega, \mathcal{F}, P)$ be an arbitrary probability space and let A be any event. Then, the function X defined by $X(\omega) = 1$, if $\omega \in A$; $X(\omega) = 0$, if $\omega \neq A$ is known as the indicator of A, and is written as I_A. Thus

$$I_A = 1, \text{ if } A \text{ occurs} ; I_A = 0, \text{ if } A \text{ fails to occur.}$$

To show that I_A is a random variable :

If $x < 0$, $\{\omega : I_A \leq x\} = \varnothing \in \mathcal{F}$.

If $0 \leq x < 1$, $\{\omega : I_A(\omega) \leq x\} = \overline{A} \in \mathcal{F}$. If $x \geq 1$, $\{\omega : I_A(\omega) \leq x\} = \Omega \in \mathcal{F}$.

If follows that I_A is a random variable.

Obviously, $P(I_A = 1) = P\{\omega : I_A(\omega) = 1\} = P(A)$, $P(I_A = 0) = P(\overline{A}) = 1 - P(A)$.

If $A_1, A_2, \ldots, A_n$ are any events ; then $X = I_{A_1} + I_{A_2} + \ldots + I_{A_n}$ $\quad (\omega \in \Omega)$

counts the number of $A_1, A_2, \ldots, A_n$ which occur.

If A_k, $1 \leq k \leq n$ are disjoint and exclusive (i.e. form a Partition) the $X = \Sigma_k I_{A_k}$ yields the *index* of that A_i which occurs.

Obviously, this sum provides $\{X = k\} = \{I_{A_k} = 1\}$.

Definition. Simple Random Variable. A quantity X is called a simple r.v. of it can be expressed as a linear combination of a *finite* collection of indicators, i.e. if there exist events $A_1, \ldots, A_n$ and scalars $a_1, \ldots, a_n$ such that $X = \Sigma a_i I_{A_i}$, $1 \leq i \leq n$.

This representation may not be unique.

Some Properties of Indicator Random Variables

1. $I_\Omega = 1$, $I_\varnothing = 0$. 2. $I_{A'} = 1 - I_A$. 3. $I_{AB} = I_A I_B$.
4. $I_{(A \cup B)} = 1 - (1 - I_A)(1 - I_B) = I_A + I_B - I_{AB}$. 5. $I_{(A \cup B)'} = (1 - I_A)(1 - I_B)$.

Proofs. 1. $I_\Omega(\omega) = 1$, because all points $\omega \in \Omega$. $I_\varnothing(\omega) = 0$, because no points $\omega \in \varnothing$.

2. $1 - I_A = 1$, iff $\omega \in A$.

3. $I_{AB}(\omega) = 1$, iff $\omega \in A_B$, i.e. iff $I_A(\omega) = 1$, $I_B(\omega) = 1$, i.e. iff $I_A(\omega)\, I_B(\omega) = 1$.

4. $I_{(A \cup B)}(\omega) = 1$, iff $\omega \in (A \cup B)$, i.e. ω is in at least one of the sets A and B. In particular, if $\omega \in A$, then $1 - I_A = 0$, and the R.H.S. of (4) is 1. If $\omega \notin A \cap B$, $I_A = 0 = I_B$ and the R.H.S. of (4) is zero.

5. $I_{(A \cup B)'} = I_{\bar{A}\bar{B}} = I_{\bar{A}}\, I_{\bar{B}} = (1 - I_A)(1 + I_B)$, [by (2) and (3)]

Note. If A and B are disjoint, $I_{AB} = I_\varnothing = 0$, and then $I_{A \uplus B} = I_A + I_B$.

Extension to n events in all cases is immediate. Thus $I_{A_1 \cup A_2 \cup \ldots} = I_{A_1} + I_{A_2} + \ldots$ etc.

Remark. The indicator random variable is often called the *alternative* variate or *zero-one* variate or Bernoulli variate.

***Example* 1.** Let I_A and I_B be the indicator functions of the independent events A and B. If the random variable X defined by $X = I_A + I_B$ satisfies the relations

$$P(X = 2) = a,\ P(X = 1) = b,\ P(X = 0) = c$$

show that at least one of the numbers a, b, c is at least $\frac{4}{9}$.

Solution. Let $P(A) = p$, $P(B) = p'$, $0 < p, p' < 1$, write $q = 1 - p$, $q' = 1 - p'$, and consider the quadratic polynomial $P(x)$ given by

$$P(x) = (px + q)(p'x + q') = pp'x^2 + x(pq' + p'q) + qq'. \quad \ldots(1)$$

Since A and B are independent, the given relations yield

$$a = pp' \neq 0, \qquad b = pq' + p'q \neq 0, \qquad c = qq' \neq 0, \qquad (a + b + c = 1)$$

Note that, e.g. $X = 1 \Rightarrow I_A = 1, I_B = 0$ or $I_A = 0, I_B = 1$, and so on,

$$\therefore \qquad P(x) = ax^2 + bx + c. \qquad a > 0,\ b > 0,\ c > 0.$$

Observe that $x = -q/p$ and $x = -q'/p'$ yields $P(x) = 0$. [by (1)]

So $P(x) = 0$ has real roots, its discriminant must be non-negative, i.e. $b^2 - 4ac \geq 0$. Hence

$$a > 0, \quad b > 0, \quad c > 0, \quad a + b + c = 1, \quad ac \leq b^2/4. \quad \ldots(2)$$

Now, suppose that $a < \frac{4}{9}$, $b < \frac{4}{9}$, then $ac \leq b^2/4 < \frac{4}{81}$.

$$a + c = (1 - b) > (1 - \tfrac{4}{9}) \Rightarrow c > (\tfrac{5}{9} - a).$$

These inequalities yield $a[(\frac{5}{9}) - a] < (\frac{4}{81}) \Rightarrow 81a^2 - 45a + 4 = (9a - 1)(9a - 4) > 0$.

So $a < (1/9)$, since $a < (\frac{4}{9})$ by hypothesis. Thus, $c = 1 - (a + b) > 1 - (\frac{1}{9}) - (\frac{4}{9})$, i.e. $c > \frac{4}{9}$. We have, thus, proved our assertion, if $0 < p, p' < 1$. However if p or p' is zero, say $p = 0$, then $a = 0$, $b = p'$, $c = q'$. Clearly either b or c is at least $\frac{4}{9}$.

Similarly, if $p = 1$, then $a = p'$, $b = q'$, $c = 0$ and either a or b is at least $(\frac{4}{9})$.

3-13. Random Variable Criterion

A real-valued function $X(\omega)$, $\omega \in \Omega$, is a r.v. iff $\{X \le x\} \in \mathcal{F}$, for all $x \in R$.

Proof. Let X be a random variable. Then $\{X \le x\} \in \mathcal{F}$, $-\infty < x < \infty$. Now, write

$$\{X < x\} = \bigcup_{n=1}^{\infty} \{X \le x - (1/n)\}.$$

Since for all n, $\{x \le x - (1/n)\} \in \mathcal{F}$, it follows that the countable union of these sets, i.e. $\{X < x\}$ is an element of $\mathcal{F}$. Thus, $\{X < x\} \in \mathcal{F}$. Conversely, let $\{X < x\} \in \mathcal{F}$, for all x and write

$$\{X \le x\} = \bigcap_{n=1}^{\infty} \{X < x + (1/n)\}.$$

Since for all n, $\{X < x + n^{-1}\} \in \mathcal{F}$, it follows that the countable intersection of these sets, viz. $\{X \le x\}$ is an element of $\mathcal{F}$. Thus, $\{X \le x\} \in \mathcal{F}$. Hence X must be a random variable.

Theorem. If X is a random variable, then $Z = aX + b$ is also a random variable.

Proof.

$$\{(aX + b) \le x\} = \{aX \le x - b\} = \{X \le (x - b)/a\}, \text{ if } a > 0.$$

$$\{(aX + b) \le x\} = \{X \ge (x - b)/a\} = \{X < (x - b)/a\}^c, \text{ if } a < 0.$$

$$\{(aX + b) \le x\} = \{b \le x\} = \begin{cases} \Omega & \text{if } b \le x, a = 0 \\ \varnothing & \text{if } b > x, a = 0 \end{cases}$$

Thus, for all cases, $\{(aX + b) \le x\} \in \mathcal{F}$, and consequently $aX + b$ is also a r.v.

Exercise 3(a)

1. Let X be a random variable. Then X^2 is also a r.v.

   [**Aid** : $\{X^2 \le x\} = \phi$ if $x < 0$, $\{X^2 \le x\} = \{-\sqrt{x} \le x \le X \le \sqrt{x}\}$, if $x \ge 0$, $\therefore$ $\{X^2 \le x\} \in \mathcal{F}$].

2. Let X be a r.v. then $Y = 1/X$ is also a r.v. provided $P(X = 0) = 0$.

   [**Aid** :
   $$\{1/X \le x\} = \{X^{-1} \le x, X < 0\} + \{X^{-1} \le x, X > 0\} + \{X^{-1} \le x, X = 0\}$$
   $$= \{Xx \le 1\} \cap \{X < 0\} + \{xX \ge 1\} \cap \{X > 0\}, \text{ etc.}]$$

3. If X is a random variable, then so are $|X|$ and aX, (a = constant).

4. If X and Y are two random variables on a probability space $(\Omega, \mathcal{F}, P)$, then $X + Y$ is also a random variable.

   [**Aid** : Show that $\{X + Y < z\} = \bigcap [\{X < r\} \cap \{Y < z - r\}]$, where union is taken over the set of all rational numbers.]

5. Let X_i $1 \le i \le n$ be r.v.s. on a probability space $(\Omega, \mathcal{F}, P)$. Define $M_n = \max\{X_1, ..., X_n\}$ by $M_n(\omega) = \max\{X_1(\omega), ..., X_n(\omega)\}$, for all $\omega \in \Omega$, and $N_n = \min\{X_1, ..., X_n\}$ by $N_n(\omega) = \min\{X_1(\omega)\}, ..., X_n(\omega)\}$, for all $\omega \in \Omega$. Show that both M_n and N_n are random variables.

   [**Aid :** $\{M_n \le x\} = \{X_1 \le x, ..., X_n \le n\} = \bigcup_{i=1}^{n} \{X_i \le x\} \in F$. Also note that,

   $$\min\{X_1, ..., X_n\} = -\max\{-X_1, -X_2, ..., -X_n\}].$$

3-20. Distribution Function and its Properties

Let X be a one-dimensional random variable. The function F defined for all real t, by the equation $F_X(t) = P(X \le t)$ is called the *(cumulative) distribution function* of X.

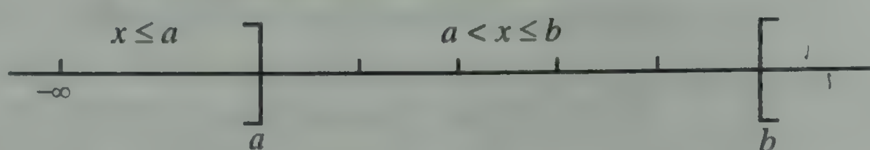
Notes. (i) We write c.d.f. for cumulative distribution function. Sometimes, only d.f. is written instead if c.d.f.

(ii) Suffix X in F_X is used to emphasize the fact that the distribution function is associated with the particular variate X. When the particular underlying variate is clear from the context, we shall simply write $F(t)$ instead of $F_X(t)$.

(iii) *Tail events.* Let x be any real number. The events $\{X < x\}$, $\{X \le x\}$, $\{X > x\}$, $\{X \ge x\}$ are called *tail events*. For distinction, we may label them *open, closed, upper* and *lower* tails. Often, simple r.v.'s. are expressed as linear combination of tail events.

Some Properties of Distribution Functions

Property 1. $\boxed{P(a < X \le b) = F(b) - F(a),}$ $(b > a)$. ...(1)



Proof. $(X \le a) \uplus (a < X \le b) = (X \le b)$ [by Point-set algebra]

$\therefore$ $P(a < X \le b) + P(X \le a) = P(X \le b)$

or $P(a < X \le b) = P(X \le b) - P(X \le a) = F(b) - F(a)$ [by transfer and Def.]

***Cors.* 1.** (i) $P(a \le X \le b) = F(b) - F(a) + p(a)$ $[P(X = c) \equiv p(c)]$

Proof. $(a \le X \le b) = (a < X \le b) \uplus (X = a)$ [by Point-set algebra]

$\therefore$ $P(a \le X \le b) = P(a < X \le b) + P(X = a) = F(b) - F(a) + P(X = a)$ [by (1)]

(ii) $P(a < X < b) = F(b) - F(a) - P(b).$

Proof. $(a < X < b) = \{a < X \le b)\} - (X = b)$ [by Point-set algebra]

$\therefore$ $P(a < X < b) = P(a < X \le b) - P(X = b) = F(b) - F(a) - P(X = b),$ $[\{X = b) \subset (a < X \le b)$ and by (1)]

(iii) $P(a \le X < b) = F(b) - F(a) + P(X = a) - P(X = b).$

Proof. $(a \le X < b) = \{a < X \le b)\} \uplus (X = a) - (X = b).$ [by Point-set algebra]

Since $\{X = b\} \subset \{a < X \le b\} \uplus \{X = a\}$, hence subtraction property of P gives,

$$P(a \le X < b) = P\{(a < X \le b) \uplus (X = a)\} - P(X = b) = P(a < X \le b) + P(X = a) - P(X = b)$$

$$= F(b) - F(a) + p(a) - p(b) \qquad \text{[by Property 1]}$$

Comments. When $p(a) = 0 = p(b)$ which is indeed the case when X is a continuous variate [see Property 4] then

$$\boxed{P(a < X \le b) = P(a \le X \le b) = P(a < X < b) = P(a \le X \le b).}$$

Property 2. (a) $0 \le F(x) \le 1 \ \forall \ x \in R$ [F is a bounded function]

(b) $F(x) \le F(y)$, if $x < y$. [F is monotonically non-decreasing]

Proof. (a) Since P is a probability measure : $0 \le P\{X \le x\} \le 1 \Rightarrow 0 \le F(x) \le 1$.

(b) $F(y) - F(x) = P(x < X \le y) \ge 0 \Rightarrow F(x) \le F(y)$. [Property 1 and $P(\cdot) \ge 0$]

Limit Property 3. (a) $L_1 = \lim\limits_{n \to -\infty} F(x) = 0$, i.e. $F(-\infty) = 0$, formally,

(b) $L_2 = \lim\limits_{n \to \infty} F(x) = 1$, i.e. $F(\infty) = 1$, formally.

Proof. (a) Let $\langle x_n \rangle$ be any decreasing sequence such that $\lim x_n \to -\infty$. Then, the sequence of decreasing intervals, viz. $\{]-\infty, x_n]\}$ yields : $\lim\limits_{n \to \infty}]-\infty, x_n] = \bigcap\limits_{n=1}^{\infty}\{]-\infty, x_n]\} = \varnothing$.

Hence, $P\left(\lim\limits_{n \to \infty}]-\infty, x_n]\right) = P(\varnothing) = 0$, i.e. $\lim\limits_{n \to \infty} P(]-\infty, x_n]) = 0$, [by Continuity of P].

Thus $\lim\limits_{n \to \infty} F(x_n) = 0 \Rightarrow F\left(\lim\limits_{n \to \infty} x_n\right) = F(-\infty) = 0.$

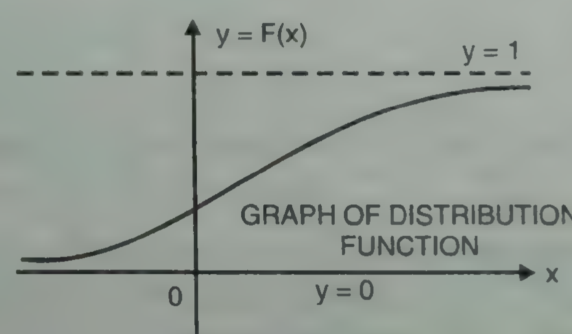
(b) Let $\langle x_n \rangle$ be an increasing sequence of real numbers such that $\lim x_n \to \infty$. Then, the sequence of increasing intervals, viz. $\{]-\infty, x_n]\}$ yields :

$$\lim_{n \to \infty}]-\infty, x_n] = \bigcup_{n=1}^{\infty}\{]-\infty, x_n]\} =]-\infty, \infty[= \Omega.$$

$$\therefore \quad P\left(\lim_{n \to \infty}]-\infty, x_n]\right) = P(\Omega) = 1, \text{ i.e., } \lim_{n \to \infty} P(]-\infty, x_n]) = 1. \quad \text{[by Continuity of } P\text{]}$$

$$\text{Thus} \quad \lim_{n \to \infty} F(x_n) = 1 \Rightarrow F\left(\lim_{n \to \infty} x_n\right) = F(+\infty) = 1.$$

These properties show that the graph of the function $y = F(x)$ must approach the line $y = 0$ at the lower end and line $y = 1$ at the upper end. In other words, the lines $y = 0$ and $y = 1$ are asymptotes to the curve $y = F(x)$.



GRAPH OF DISTRIBUTION FUNCTION

Property 4. (i) $\lim\limits_{t \to a^+} F(t) = F(a)$,

(ii) $\lim\limits_{t \to a^-} F(t) = P(X < a) = F(a) - P(X = a)$.

These limit relations reveal that :

(i) A distribution function is continuous from the right.

(ii) F has a jump discontinuity at $t = a$ when $t \to a$ from the left.

Proof. (i) Let $\langle x_n \rangle$ be a decreasing sequence of real numbers such that $x_n > a$ and $\lim x_n = a$. Then the sequence of intervals, viz. $([-\infty, x_n]\}$ is a decreasing sequence of intervals and so

$$\lim_{n \to \infty}]-\infty, x_n] = \bigcap_{n=1}^{\infty}\{]-\infty, x_n]\} =]-\infty, a].$$

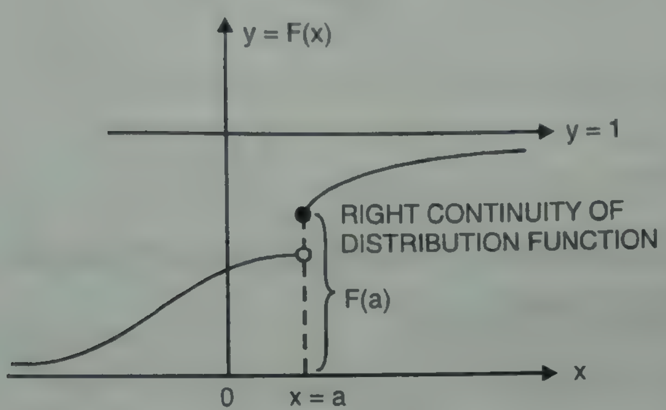
$$\therefore \quad P\{\lim_{n \to \infty}]-\infty, x_n]\} = P(]-\infty, a]) = P(X \le a) = F(a),$$

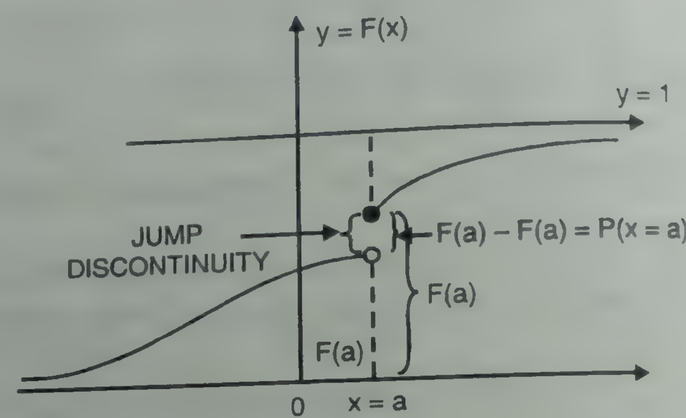
$$\text{i.e.} \quad \lim_{n \to \infty} P\{]-\infty, x_n]\} = F(a), \text{ [} P \text{ is continuous].}$$

Thus $\lim_{n\to\infty} F(x_n) = F(a)$, or $F(a+) = F(a) \Rightarrow F$ is continuous from the right.

(ii) Let $\langle x_n \rangle$ be an increasing sequence of real numbers such that $x_n < a$, $\forall\, n$ and $\lim x_n = a$. Then the sequence of interval $\{]-\infty, x_n]\}$ is an increasing sequence of intervals and therefore,

$$\lim_{n\to\infty}]-\infty, x_n] = \bigcup_{n=1}^{\infty} \{]-\infty, x_n]\} =]-\infty, a].$$





$$\therefore\quad P\left\{\lim_{n\to\infty}]-\infty, x_n\right\} = P\{]-\infty, a]\} \text{ or } \lim_{n\to\infty} P(]-\infty, x_n]) = P(X < a), \text{ [by Continuity of } P]$$

i.e. $\lim_{n\to\infty} F(x_n) \quad = P(X < a) \Rightarrow F(a-) = P(X < a)$

As $\{X \le a\} = (X < a) \uplus (X = a)$, hence

$$P(X \le a) = P(X < a) + P(X = a) \Rightarrow F(a-) \equiv P(X < a) = F(a) - P(X = a).$$

Thus F has a jump discontinuity at $x = a$ where $x \to a$ from the left, with the jump value :

$$P(X = a) = F(a) - F(a-).$$

Remark. If c.d.f. F is continuous at a, then $F(a) = F(a-)$ and so $P(X = a) = 0$, i.e. if the c.d.f. is continuous, then the underlying random variable takes any particular (point) value with probability zero.

Summary of Some Properties of a c.d.f.

1. $0 \le F(x) \le 1, \quad \forall\, x \in R.$ [Boundedness Property]
2. F is non-decreasing, i.e. if $x \le y$, then $F(x) \le F(y)$. [Monotone increasing Property]
3. $F(-\infty) = 0$, $F(+\infty) = 1$. That is $\lim_{n\to-\infty} F(x_n) = 0$, $\lim_{n\to\infty} F(x_n) = +1$. [Limits Property]
4. F is continuous from the right at each point, i.e. $F(a+) = F(a)$. [Right-continuous Property]
   $F(a+) - F(a-) = P(X = a)$. [Jump Discontinuity]
5. $P\{a < X \le b\} = F(a) - F(a)$. [Interval Property]

[Conditions (2), (3) and (4) are necessary as well as sufficient for F to be a c.d.f. on R]

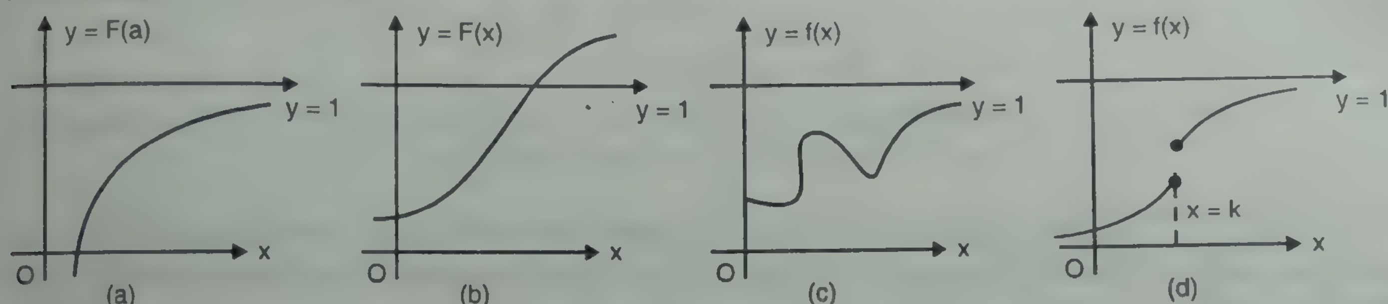
Note. $\lim\limits_{x \downarrow a}$ means the limit is taken with x decreasing to a (i.e. from the right) and $\lim\limits_{x \uparrow a}$ means the limit is taken with x increasing to a (i.e. from the left). We often write $\lim\limits_{x \downarrow a} F(x) = F(a+)$, $\lim\limits_{x \uparrow a} F(x) = F(a-)$.

Comments. If we define $F(x) = P\{X < x\}$, then it is left continuous, not right continuous, vide

$$P(X = \pm 1) = \tfrac{1}{4},\ P(X = 0) = \tfrac{1}{2}.$$

3-21. *Worked-out Problems*

Example 1. Give reasons why each of the graphs of F given below does not represent a distribution function.



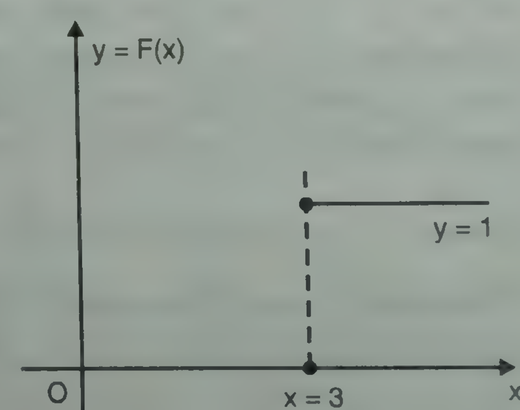
Solution. (a) $F(x) < 0$ (negative) for some x. (b) $F(x) > 1$ for some x.
(c) F is not non-decreasing i.e. sometimes F is decreasing also.
(d) F is not right continuous at $x = k$, in fact it is left continuous.

Example 2. Show that the function F defined by :
$F(x) = 0, x \le 3$; $F(x) = 1, x > 3$ is not distribution function.

Solution. Here $0 \le F(x) \le 1\ \forall\ x \in R$, and $F(-\infty) = 0$, $F(+\infty) = 1$. Also F is monotonically increasing. However,

$$F(3+) = \lim_{x \to 3+} F(x) = 1;\quad F(3) = 0, \text{ since } F(x) = 1\ \forall\ x > 3.$$

As $F(3+) \ne F(3)$, F is not right-continuous, so F is not a c.d.f.



Problems with Solutions Provided at the End of the Text

1*. Show that the function F defined by,
$F(x) = 0, x = 0$; $F(x) = k \ln x, 0, \le x < 3$; $F(x) = 1, x \ge 3$
cannot represent a c.d.f. for any non-zero constant k.

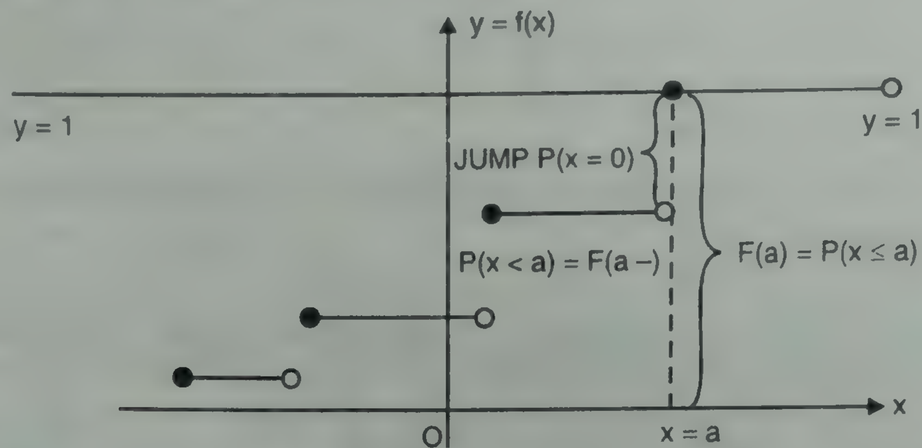
2*. (i) Show that, if $a \le X \le b$, then $F(x) = 0$ for $x < a$, and $F(x) = 1$ for $x > b$.
(ii) Show that $X(\omega) \le Y(\omega)\ \ \forall\ \omega \in \Omega \Rightarrow F_Y(t) \le F_X(t)$.

3-30. *Discrete Distribution*

Let X be a random variable defined on an arbitrary probability space $(\Omega, \mathcal{F}, P)$. The distribution function F defined by $F(x) = P(X \le x)$ plays a pivotal role in the theory of probability because it completely determines the probability distribution of X. In § 3-21(4) we noted that if F is continuous at a, then $P(X = a) = 0$. And if F is not

continuous at a, then $P(X = a) = F(a) - F(a-) > 0$ is the jump in F at a. Accordingly, we categorise random variables on the basis of their distribution functions.

***Definition* 1.** Let X be a r.v. on an arbitrary probability space $(\Omega, \mathcal{F}, P)$. If the c.d.f. F of X is discontinuous and its graph is a *stair case* form (shown in the Fig.), then X is said to be a *discrete* r.v.



Since the d.f F of any variate X is bounded and monotonically increasing, it (F) can have at most countable number of points of (jump) discontinuities. Label these points of discontinuities as $x_1, x_2, ..., x_n$ and observe that :

$$P(X = x_i) = F(x_i) - F(x_i-) > 0, \ \forall\ i\ ;\ P(X = x) = 0\ \ \forall\ \ x \notin \{x_1, x_2, ...\}.$$

Let $p(x) = P(X = x)$, then p is a non-negative function which is positive only at the points $x_1, x_2, x_3, ...$ This leads to an alternate definition of a discrete r.v.

***Definition* 2.** A r.v. X is said to be discrete, if there exists a non-negative function p which vanishes everywhere except at a finite or at a countably infinite number of points such that its c.d.f. is given by

$$F(x) = \sum_{x_i \le x} p(x_i),\ \forall\ x.$$

Probability Mass Functions. The function p given by $p(x) = P(X = x) \equiv f(x)$, say is called *probability mass function* (p.m.f..) or *discrete density* of a (discrete) variate X. Observe that : $P\{\omega : X(\omega) \in A\} \equiv P(X \in A) = P(A) = \Sigma p(x),\ \ \forall\ x \in A.$

3-31. *Discrete Density Theorem*

Let X be a discrete r.v. with p.m.f. $p(x)$; then (a) $p(x) \ge 0$, (b) $\sum_{i=1}^{\infty} p(x_i) = 1.$

Proof. (a) Since $p(x)$ is the probability function (p.f.) of X, so $p(x) = P(x = x) \ge 0,\ \forall\ x.$

(b) By definition, $F(x) = \sum_{x_i \le x} p(x_i),\ \forall\ x$, so $\lim_{n\to\infty} F(x) = \sum_{x_i < \infty} p(x_i)$. ...(1)

Since $\lim_{x\to\infty} F(x) = F(+\infty) = 1,\ \sum_{x_i \le \infty} p(x_i) = \sum_{i=1}^{\infty} p(x_i)$ result (1) provides the statement (b).

3-32. Worked-out Problems

***Example* 1.** If the variate X denotes the maximum score obtained in k independent tosses of a fair Die, find the p.m.f.. of X.

Solution. Let $X_1, X_2, ..., X_k$ denote the k-scores, so that $X = \max\{x_1, x_2, ..., x_k\}$. Now
$P\{X \le x\} = P\{\max(X_1, X_2, ..., X_k) \le x\} = P\{X_1 \le x, X_2, \le x, ..., X_k \le x\}$
$= P(X_1 \le x) \,.\, P(X_1 \le x) \ldots P(X_k \le x), = \{P(X_i \le x)\}^k = (x/6)^k$. [By indep. and by symmetry]
$\therefore\ P(X = x) = P(X \le x) - P(X \le x - 1) = (x/6)^k - [(x-1)/6]^k = [x^k - (x-1)^k]/6^k,\ 1 \le x \le 6$ ...(1)
For 2 dice $k = 2$, $P(x = x) = (2x - 1)/36,\ 1 \le x \le 6$.
Note that as $k \to \infty$.

$$P(X = x) = \lim_{k\to\infty}\left(\frac{x}{6}\right)^k\left[1-\left(1-\frac{1}{x}\right)^k\right] = \begin{cases}1, & \text{if } x = 6\\ 0, & \text{if } x < 6\end{cases}.$$

***Example* 2.** A random variable X has the following probability distribution :

x	0	1	2	3	4	5	6	7	8
$p(x)$	k	$3k$	$5k$	$7k$	$9k$	$11k$	$13k$	$15k$	$17k$

(a) Determine the value of k. (b) Find $P(X < 4)$, $P(X \ge 5)$, $P(0 < X < 4)$.
(c) Find the c.d.f. (d) Find the smallest value of x for which $P(X \le x) > 0.5$.

Solution. (a) $\Sigma p(x_i) = 1 \Rightarrow k(1 + 3 + 5 + 7 + ... + 17) = 1 \Rightarrow k = 1/81$.

(b)
$$P(X < 4) = P\{(X = 0) \uplus (X = 1) \uplus (X = 2) \uplus (X = 3)\}$$
$$= p_0 + p_1 + p_2 + p_3 = k + 3k + 5k + 7k = 16k = 16/81.$$
$$P(X \ge 5) = P\{(X = 5) \uplus (X = 6) \uplus \ldots \uplus (X = 8)\} = p_5 + p_6 + p_7 + p_8 = 56k = 56/81.$$
$$P\{0 < X < 4\} = P\{(X = 1) \uplus (X = 2) \uplus (X = 3)\} = p_1 + p_2 + p_3 = 15k = 15/81.$$

(c) The c.d.f. is $F(t) = P(X \le t)$.

$F(0) = P(X \le 0) = k,$ $F(1) = P(X \le 1) = 4k$ $F(2) = P(X \le 2) = 9k$
$F(3) = P(X \le 3) = 16k,$ $F(4) = P(X \le 4) = 25k,$ $F(5) = P(X \le 5) = 36k$
$F(6) = P(X \le 6) = 49k,$ $F(7) = P(X \le 7) = 64k,$ $F(8) = P(X \le 8) = 81k.$

Observe that $P(X < 4) = P(X \le 3) = F(3) = 16k$, etc.
$$P(X \ge 5) = 1 - P(X < 5) = 1 - P(X \le 4) = 1 - 25k = 86/81.$$
$$P(0 < X < 4) = F(4) - F(0) - p_4 = (25 - 1 - 9)k = 15k = 15/81.$$

(d)
$$F(x) = \tfrac{1}{2};\ F(5) = 36/81 = 0.44;\ F(6) = 49/81 = 0.61.$$

Thus, the smallest value of x for which $F(x) > \frac{1}{2}$ is $x = 6$.

Indicator Notation. It is usual to express $F(x)$ in expanded form or in condensed form using indicators for individual ranges.

$$F(x) = \begin{cases} 0 & -\infty < x < 0\\ k & x \le 0\\ 4k & x \le 1\\ 9k & x \le 2\\ \ldots\ldots & \end{cases}$$

$F(x) = 0\, I_{(-\infty, 0)}(x) + k\, I_{(0, 0]}(x) + 4k\, I_{(0, 1]}(x) + 9k\, I_{(0, 2]}(x) + 16k\, I_{(0, 3]}(x) = 25k\, I_{(0, 4]}(x)$.

We may even rewrite the indicator form as

$$F((x) = 0 \;\; I(-\infty < x < 0) + k\, I(0 < x \le 0) + 4k\, I(0 < x \le 1) + 9k\, I(0 < x = 2) + \ldots$$

Problems with Solutions Provided at the End of the Text

1*. A random variable X takes values 0, 1, 2, 3, ... with probability proportional to $(x + 1)\,(1/5)^x$. Find $P\{X \le 5\}$.

2*. The variate X has p.m.f. $p(x) = 2(x + 1)/(n + 1)\,(n + 2)$, $x = 0, 1, 2, \ldots, n$. Events A and B are defined by $\{X \le m\}$ and $\{X \le \ell\}$ respectively, where $0 \le \ell \le m < n$. Construct an appropriate model conditioned on A and also evaluate $P\{B \mid A\}$.

3*. A weighted coin, $[P(h) = \frac{2}{3}, P(t) = \frac{1}{3}]$ is tossed three times. If the variate X denotes the number of heads produced in three tosses, find the c.d.f $F_X(x)$ and p.m.f. $f(x)$.

4*. A random sample of three teachers is chosen from the list of Academic Council (A.C.) members of Delhi University. If the proportion of A.C. members who are senior teachers is 0.4 and the number of A.C. members is large, find the p.m.f. and the distribution function for the number of senior teachers in the sample.

5*. The game of chuck-a-luck is played as follows : Three fair dice are rolled. A player A bets Rs. 1, on the occurrence of one of the integers 1, 2, 3, ..., 6, say k. Then if one k occurs (on the 3 dice) A wins Rs. 1, if two k's occur A wins Rs. 2, and if three k's occur, A wins Rs. 3. If no k's occur, A loses Rs. 1. If variate X denotes the net amout A wins in one play of his game, find the distribution function for X and compute $P(0 < X \le 3)$, $P(X \le 0)$, $P(-1 < X \le 0)$.

6*. Let $P\{X = a/b\} = K(e^{a+b} - 1)^{-1}$, $(a, b) = 1$, (a, b +ve integers).
Find K and evaluate $P(X < 1)$ and $P(X \le 1)$.

3-40. *Some Standard Discrete Distributions*

Some commonly occuring probability mass functions (p.m.f.) have acquired definite terminology. We quote these probability laws (distributions) which will from a part of solid core in the ongoing study of Probability Theory.

1. ***Degenerate Distribution* :**

Let D be a *discrete* subset of R and $c \in D$. If $P\{X = c\} = 1$ and $P\{X \ne c\} = 0$, then $P : D \to R$ is called *degenerate distribution*. It occurs naturally as the distribution of a *constant* r.v. viz. $X : \Omega \to R$ such that $X(\omega) = c,\ \forall\ \omega \in \Omega$.

2. ***Uniform Distribution* :**

Let S be a *finite* subset of R, say $S = \{x_1, x_2, \ldots, x_n\}$. If

$$P(X = x_i) \;\equiv f(x_i) = 1/n,\ i = 1, 2, \ldots, n\ ;\ f(x_i) = 0, \text{ otherwise},$$

then f is called a *uniform distribution*. It is an index of equi-likelihood in probabilistic modelling.

Notation : $X \sim U\{x_1, x_2, \ldots, x_n\}$ or $X \sim$ unif $(x_1, x_2, \ldots, x_n)$.

3. ***Bernoulli Distribution*** :

Let $0 < p < 1$, $p + q = 1$, and define $P(X = 1) \equiv f(1) = p$, $P(X = 0) \equiv f(0) = q$. The f is called Bernoulli distribution. It deals with indicator r.v., *viz*. $X_A(\omega) = 1$, $\omega \in A$; $X_A(\omega) = 0$, $\omega \notin A$; such that $P(X_A = 1) = P(A) = 1$; $P(X_A = 0)$ $P(\bar{A}) = q$. We shall often write $f(x) = P(X = x) = q^{1-x} p^x$, $x = 0, 1$. ***Notation*** : $X \sim \text{Ber}(p)$, or bin $(1, p)$.

4. ***(Bernoulli) Binomial Distribution*** :

Let $0 < p < 1$, $p + q = 1$, $n \geq 1$ and define

$$P(X = x) \equiv f(x) = \binom{n}{x} p^x q^{n-x}, x = 0, 1, 2, \ldots, n. \quad f(x) = 0, \text{ otherwise.}$$

Then f is called a *binomial distribution* with parameters n and p. *Notation* : $X \sim$ bin (n, p). Here r.v. X records the number of successes in n Bemoulli trials with constant success-probability p.

Note. $f(x) = \sum_{k=0}^{n} \binom{n}{k} q^{n-k} p^k \,\delta(x-k)$; $\quad F(x) = \sum_{k=0}^{n} \binom{n}{k} q^{n-k} p^k \,u(x-k)$.

5. ***Poisson Distribution*** :

Let $P(X = x) \equiv f(x) = \dfrac{e^{-\lambda}\lambda^x}{x!}, 0, 1, 2, \ldots, \infty$; $\quad f(x) = 0$, otherwise.

Then f is called a *Poisson distribution* with parameter λ (> 0). [***Notation*** : $X \sim$ Pois (λ)]

Note. $f(x) = \sum_{k=0}^{\infty} \dfrac{e^{-\lambda}\lambda^x}{k!} \delta(x-k)$; $\quad F(x) = \sum_{k=0}^{\infty} \dfrac{e^{-\lambda}\lambda^x}{k!} u(x-k)$.

6. ***Geometric Distribution*** :

$$P(X = x) \equiv f(x) = q^x p, \; x = 0, 1, 2, 3, \ldots \quad [\textit{Notation} : X \sim \text{geom}\ (p)]$$

Pascal Geometric Distribution or f.f.t. Distribution

$P(Y = y) \equiv f(y) = q^{y-1} p$, $y = 1, 2, 3, \ldots$ $\quad$ ***Notation*** : $Y \sim$ gem (p)

P[First success on trial No. $x\} = q^{y-1} . p$. $\quad$ [f.f.t. : for the first time]

Note that $X = Y + 1$

7. ***Negative Binomial Distribution*** :

$$P\{X = x\} \equiv f(x) = \binom{x+k-1}{k-1} p^k q^x \equiv \binom{-k}{x} p^k (-q)^x, x = 0, 1, 2, 3, \ldots \quad [\textit{Notation} : X \sim \text{NB}(k, p)]$$

Here X = number of failures before the kth success. Note that geom$(p) \equiv$ NB$(1, p)$.

Pascal Neg-binomial Distribution : $\quad$ [***Notation*** : $Y \sim$ NB*(k, p).]

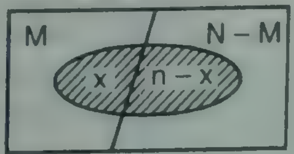
$$P(Y = y) \equiv f(y) = P\{r\text{th success on trial No. } y\} = \binom{y-1}{k-1} p^k q^{y-k}, y = k, k+1, k+2, \ldots$$

Note that $Y = X + k$ and gem (p) = NB*$(1, p)$.

$$\left[\sum_{y=k}^{\infty} f(y) = p^k \sum_{y=k}^{\infty} \binom{y-1}{k-1} q^{y-k} = p^k \sum_{x=0}^{\infty} \binom{k+x-1}{k-1} q^x = p^k (1-q)^{-k} = 1\right] \quad \left[\binom{n}{r} = \binom{n}{n-r}\right]$$

8. ***Hypergeometric Distribution*** :

$$P(X = x) = \binom{M}{x}\binom{N-M}{n-x} \Big/ \binom{N}{x},$$



H-G Model

$\max\{0, M+n-N\} \le x \le \min(M, n)\}$ **Notation :** $X - HG(N, M, n)$.

9. ***Logarithmic Series Distribution :***

$$f(x) = -\theta^x/[x \ln(1-\theta)],\ 0 < \theta < 1, \quad x = 1, 2, \ldots$$

10. ***Multinomial Distribution :*** $X = (X_1, X_2, \ldots, X_k)$

$$P\{X_1 = x_1, X_2 = x_2, \ldots, X_k = x_k\} = \binom{n}{x_1, x_2, \ldots, x_k} p_1^{x_1} \cdot p_2^{x_2} \cdots p_k^{x_k}$$

where : $0 < p_i < 1$, $p_1 + p_2 + \ldots + p_k = 1$, $x_1 + x_2 + \ldots + x_k = n$. **Notation.** $X \sim \text{Mul}(n, k\,;\, p_i)$

Trinomial Distribution :

$$f(x, y) = P(X = x, Y = y) = \binom{n}{x, y, n-x-y} p^x q^y (1-p-q)^{n-x-y}.$$

$x, y = 0, 1, 2, \ldots, n$, $x + y \le n$, $0 \le p, q < 1$, $p + q \le 1$.

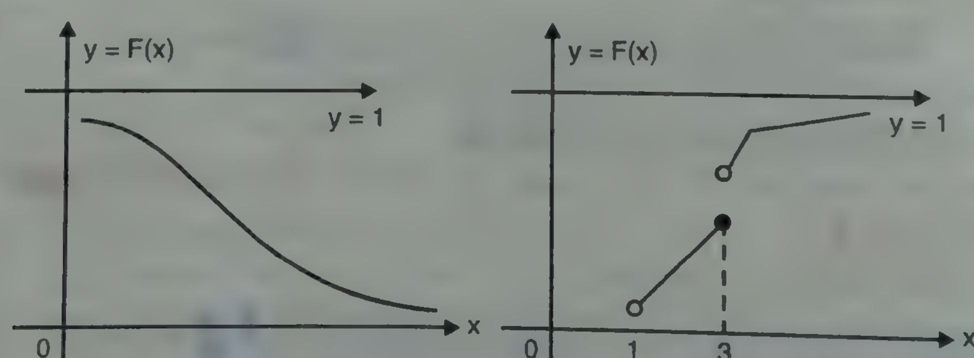
Another Usage : $f(x_1, x_2, x_3) = \binom{n}{x_1, x_2, x_3} p_1^{x_1} p_2^{x_1} p_3^{x_1}$, $x_1 + x_2 + x_3 = n$, $p_1 + p_2 + p_3 = 1$,

$0 \le p_1, p_2, p_3 \le 1$.

Comment. In all distributions, we always take $P(X = x) = 0$, when x is different from the stated range (or support) of X.

Exercise 3(b)

1. Give reasons why each of the graphs of F given below does not represent a distribution function.



2. Show why a function F defined as under, does not represent a c.d.f. ?

$$F(x) = \begin{cases} 0 & x \le 2 \\ \frac{1}{3} & 2 < x < 3 \quad [F(2+0) \ne F(2)] \\ 1 & x \ge 3 \end{cases}$$

3. If a variate X assume three values 0, 1, 2 with probabilities $\frac{1}{3}, \frac{1}{6}, \frac{1}{2}$ respectively, find the c.d.f. of X and show that $P(X \le 1) = \frac{1}{2}$.

4. Let $\Omega = \{(a, b), : a, b \in \{1, 2, \ldots, 6)\}$ and $\mathcal{F}$ be the set of all sub-sets of Ω. Let $P\{(a, b)\} = 1/36$, for all 6^2 pairs (a, b) in Ω. Define $X(a, b) = a + b$, $1 \le a, b, \le 6$. Obtain the c.d.f. of X.

5. A random variable X can take all non-negative integral values and the probability that X takes the value r is proportional to α^r, $(0 < \alpha < 1)$. Prove that $P(X = c) = (1-\alpha)\alpha^c$.

6. A random variable X assumes the values 1, 2, 3, 4, 5 such that

$$P(X=1)=P(X=2)\ ;\ P(X=4)=P(X=5)\ ;\ P(X<3)=P(X=3)=P(X>3).$$

Write down the distribution of X and show that $P(X \le 3) = \frac{2}{3}$.

7. A random variable X assume the values $-3, -2, -1, 0, 1, 2, 3$ such that

$P(X>0)=P(X=0)=P(X=0),\ P(X<-3)=P(X=-2)=P(X=-1),\ P(X=1)=P(X=2)=P(X=3).$

Write down the distribution of X and show that $P(X \le 1) = \frac{7}{9}$.

8. Let $p(x)$ be the probability function of a discrete variable X which assume the values x_1, x_2, x_3, x_4 such that $2p(x_1)=3p(x_2)=p(x_3)=5p(x_4)$.
Show that the p.m.f.. of X is $15\lambda, 10\lambda, 30\lambda, 6\lambda$, where $\lambda = 1/31$.

9. Let $p(x) = x/5$, $x = 1, 2, 3, 4, 5$, $p(x) = 0$, elsewhere. Show that $P(X = 1 \text{ or } 2) = \frac{1}{5}$ and $P\{\frac{1}{2} < X < (\frac{5}{2}) \mid X > 1\} = \frac{1}{7}$.

$x/15$

10. The p.m.f. $p(x)$ of variate X has values.

$$p(0)=3k^3,\ p(1)=4k-10k^2,\ p(2)=5k-1,\ p(x)=0 \text{ elsewhere, } k>0.$$

(a) Find k and $P(X<1)$, $P(X<2)$, $P(1<X\le 2)$, $P(0<X<3)$.

(b) Describe the c.d.f. $F(\cdot)$ of X and sketch its graph.

(c) Find the largest a such that $F(a) < 1/2$.

(d) Find the smallest b such that $F(b) > 1/3$.

[**Ans.** $k=\frac{1}{3}$, $P(X<2)=\frac{1}{3}$, $P(0<X<3)=\frac{8}{9}$, $a=1, b=2$]

11. Let X be a random variable with p.m.f.. given by

x :	-2	-1	0	1	2	3	7	12
$f(x)$:	0.1	0.15	0.2	0.2	0.15	0.1	0.05	0.05

What is the probability of the event that :

(a) X is even ? (b) X is negative ? (c) X is less than (-1) and odd ? (d) X takes a value between 1 and 8 inclusive (e) $X^2 < 9$? [**Ans.** 0.5 ; 0.25 ; 0 ; 0.5 ; 0.8]

12. A variable X has the following p.m.f. :

x :	0	1	2	3	4	5	6	7
$f(x)$:	0	k	$2k$	$2k$	$3k$	k^2	$2k^2$	$7k^2+k$

(a) Find k and evaluate $P(X>4)$, $P(X<6)$, $P(X^2<6)$ and $P(1<X^2<17)$,

(b) Determine the c.d.f. of X and $P(X \le 5)$.

(c) Find the smallest value of x for which $P(X \le x) > 1/2$.

[**Ans.** (a) $k = 1/10, 0.81, 0.30, 0.70$, (c) min $x = 4$]

13. For a variate X; $P\{X=k+1\} = \alpha P\{X=k\}$; $k = 0, 1, 2, 3, \ldots, 0<\alpha<1$.
Find the p.m.f. for X and show that $P(X = \text{odd number}) = \alpha/(1+\alpha)$.

14. A variate X ranges over : $R_x = \{-n, -n+1, \ldots, 0, \ldots, 1, 2, \ldots, n+1\}$.
Suppose we want

$$P(X=-k) = P(X=k) \text{ and } P\{|X|=k+1\} = \tfrac{2}{3}P\{|X|=k\}.$$

Find the p.m.f. for X and evaluate $P(X \ne 0)$ and $P(|X| = 3j)$ for some integer j.

15. A random variable X assumes values which are rational numbers of the form $n/(n+1)$ and $(n+1)/n$, where $n = 1, 2, 3, \ldots$. If

$$p[X = n/(n+1)] = P[X = (n+1)/n] = \left(\tfrac{1}{2}\right)^{n+1},$$

verify that this assignment of probabilities is permissible and find the c.d.f. of X.

16. An urn contains 4 balls numbered 1, 2, 3, 4 respectively. Two balls are drawn without replacement. Let X be the sum of the two numbers that occur. Obtain the probability function for X. How is the probability function modified if the two balls are drawn with replacement ?

[**Ans.** $f(3) = f(4) = \frac{1}{2} f(5) = f(6) = f(7) = p = \frac{1}{6}$

$f(2) = \frac{1}{2} f(3) = \frac{1}{3} f(4) = \frac{1}{4} f(5) = \frac{1}{3} f(6) = \frac{1}{2} f(7) = f(8) = p = \frac{1}{16}$.]

17. Prove that : $P(A \mid X = x) = P(B \mid X = x) \;\forall\; x \le c \Rightarrow P(A \mid X \le c) = P(B \mid X \le c)$.

18. A random variable X has the following probability function :

x :	−3	−2	−1	0	1	2	3
$p(x)$:	0.08	0.14	0.19	0.27	0.17	0.09	0.06.

Find the density function of $Y = \{ \mid X - 1 \mid + \mid X + 1 \mid \}$. [**Ans.** $f(2) = 0.63, f(4) = 0.23, f(6) = 0.14$]

19. An experiment consists of three independent tosses of a fair coin. Let X, Y, Z denote the number of heads, the number of head-runs and the length of head runs respectively. [A *head-run* is defined as consecutive occurrences of at least two heads. Lengh of head-runs is the number of heads occurring together]. Find the probability function of X, Y, Z, $U = X + Y$ and $V = XY$. Also construct probability tables and draw their probability charts.

{**Ans.** $x = 0, 1, 2, 3, p_x = p, 3p, 3p, p\,; y = 0, 1, p_y = 5p, 3p$.

$z = v = 0, 1, 2, 3, p_x = p_v = 5p, 0, 2p, p, u = 0, 1, 2, 3, 4\,; p_u = p, 3p, p, 2p, p$]

20. An organism of n cells is examined to determine the number of live cells it contains. A histologist suggest that the density function describing this number is

$$p(x) = \theta^x (1 - \theta), x = 0, 1, 2, ..., n - 1\,; p(x) = \theta^n, x = n$$

where $0 < \theta < 1$. On this basis, find the probability that an organism contains both dead and live cells, and show that it is less than $1 - (1/n)$.

21. A random variable X has density function $p(x)$ = defined by

x :	1	2	3	4	5	6
$p(x)$:	0.04	0.15	0.37	0.26	0.11	0.07

What is the appropriate model conditioned on the event $A = \{X \ge 4\}$? Verify that : $P\{X = 4 \mid A\} + P\{X = 5 \mid A\} + P\{X = 6 \mid A\} = 1$.

Show that (a) $P\{X \text{ odd} \mid X < 5\} = \frac{1}{2}$, (b) $P(X < 5 \mid X \text{ odd}\} = 41/52$, (c) $P\{X = 4 \mid X \ne 3\} = 26/63$.

22. The number of operations required to complete a certain process is variable, but the process is certainly completed in at most $2n$ operations. The mechanism of the process suggests that the density function describing the number of operations required is

$$p(x) = \theta^{x-1}(1 - \theta), x = 1, 2, ..., 2n - 1\,; p(x) = \theta^{2n-1}, x = 2n$$

where $0 < \theta < 1$. What are the probabilities that :

(a) the process will be completed before the kth operation ?

(b) at least k and at most $2n - k$ operations are required to complete the process ?

If the supervisor 1 attends odd-numbered operations and supervisor 2 attends even-numbered operations, show that supervisor 1 is the more likely to see the process completed provided $2\theta^{2n} + \theta - 1 < 0$.

3-50. Continuous Random Variable

1. *Definition.* Let X be a random variable on a probability space $(\Omega, \mathcal{F}, P)$ with F as the c.d.f. of X. The variate X is said to be *continuous* if there exists a non-negative function f, $[f(x) \geq 0 \ \forall \ x \in R]$ on the real line such that

$$F(x) = \int_{-\infty}^{\infty} f(t)\, dt \qquad ...(1)$$

is a continuous function of x. Thus, r.v. X is continuous, if its c.d.f. is continuous. The function f, satisfying relation (1) is called the probability density function (p.d.f.) of X.

Remarks. The word *continuous* in the above definition is misnomer. A random variable X is a function and the notion of continuity for any function in Mathematics is well established. But in the case of continuous variate, the word *continuous* is not used in its usual sense.

Even random variable is, neither random nor variable, but is a function.

2. *Absolutely Continuous Function.* A function which may be expressed as an indefinite integral is called *absolutely continuous function.*

The following facts about absolutely continuous functions are assumed :

(i) Absolutely continuous function is always continuous, but the converse need not be true.

(ii) Absolutely continuous function is differentiable everywhere except at a finite or countable number of points.

These facts establish that a variate X is continuous iff its c.d.f. F is absolutely continuous. Hence, for a continuous variate X, the c.d.f. 'F' is always continuous. Note carefully : a continuous variate is one whose c.d.f. is continuous and further it can be obtained by integrating a non-negative function f as in (1).

Singular Distributions. There are examples of variates whose c.d.f. is just continuous but the variates themselves are not continuous, i.e. the c.d.f. is not absolutely continuous. These types of distributions are known as *singular distributions.*

Singular distributions are seldom considered in this book.

3-51. *Theorem*

Let X be a continuous variate with p.d.f. 'f'. Then,

(i) $P\{X = a\} = 0$ (Point-probability is zero)

(ii) $\int_{-\infty}^{\infty} f(x)\, dx = 1,\ f(x) \geq 0,$ (Normalization Property)

(iii) $P\{a \leq X \leq b\} = \int_a^b f(x)\, dx.$

Proof. Since X is a continuous variate with p.d.f. f, the c.d.f. F is continuous.

$$F(x) = \int_{-\infty}^{\infty} f(t)\, dt. \qquad ...(1)$$

(i) Since F is continuous, $F(a-) = F(a)$, and thus $P(X = a) = F(a) - F(a-) = 0.$

(ii) $1 = F(+\infty) = \lim_{n\to\infty} F(x) = \lim_{n\to\infty} = \int_{-\infty}^{b} f(t)\,dt = \int_{-\infty}^{a} f(t)\,dt$.

(iii) $P(a \le X \le b) = F(b) - F(a) = \int_{-\infty}^{b} f(t)\,dt - \int_{-\infty}^{a} f(t)\,dt = \int_{a}^{b} f(t)\,dt.$

Remarks. (i) For a continuous variate, point probabilities are zero.

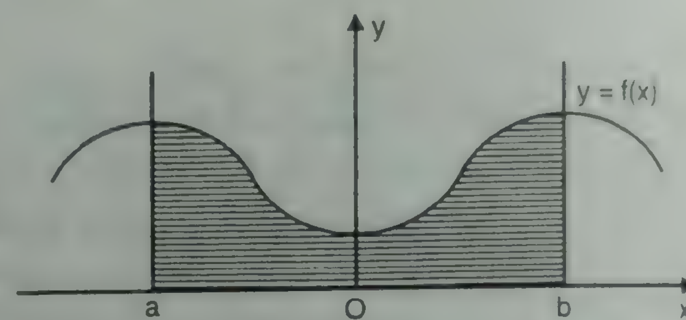
(ii) Area under the probability curve $y = f(x)$ is unity ; the fact $f(x) \ge 0$ implies that graph of $f(x)$ is above x-axis.

(iii) Area under the graph of probability curve $y = f(x)$ bounded by $x = a$, $x = b$ is simply $P(a \le X \le b)$.

3-52. *Nomenclature of p.d.f.*

The relation between p.d.f. and c.d.f. is

$$F(x) = \int_{-\infty}^{x} f(t)\,dt \qquad \text{...(1)}$$



By the Fundamental Theorem of Calculus,

$$dF/dx = F'(x) = f(x). \qquad \text{...(2)}$$

at every point where f is continuous. As per definition of derivative.

$$f(x) = \frac{d}{dx}F(x) = \lim_{h\to 0}\frac{F(x+h)-F(x)}{h} = \lim_{h\to 0}\frac{P(X \le x+h) - P(x \le h)}{h}$$

$$= \lim_{h\to 0}\frac{P(x \le X \le x+h)}{h} = \lim_{h\to 0}\frac{P\{X \in]x, x+h]\}}{h}.$$

This shows that density f represents the limit of the ratio of the probability over an interval of length, to the length h of that interval, as $h \to 0$. It is for this reason that f is called p.d.f.

Comments. The result $f(x) = F'(x)$ enables us to derive the density of a continuous variate X from its c.d.f. But the equation is valid only for the continuity points of f. We recall from the properties of absolutely continuous functions that F will be differentiable everywhere except for a finite or countably infinite number of points. The points where the c.d.f. 'F' is not differentiable are those points where the graph of F has sharp edges, and at such points we assign to f, any value we please, without disturbing the basic probability law associated with the random variable. We shall assign values to f in the most convenient way, suitable to given situation. As such, the density f and c.d.f. F are always connected by

(i) $F(x) = \int_{-\infty}^{x} f(t)\,dt \;\; \forall \;\; x \in R,$ (ii) $\dfrac{dF(x)}{dx} = f(x) \;\; \forall \;\; x \in R.$

These observations are exemplified in the following problems :

3-53. Worked-out Problems

***Example* 1.** In terms of continuous $f_X(\cdot)$ and $F_X(\cdot)$, find
(i) $P[X > a + \Delta x \mid X > a]$, (ii) $P[a < X < a + \Delta x \mid X > a]$, (iii) lim (ii)/$\Delta x$ as $\Delta x \to 0$.

Solution. For precision, write $\Delta x = h$; and observe that X is continuous Now,

$$\text{(i) } p_1 = \frac{P(X > a+h, X > a)}{P(X > a)} = \frac{(X > a+h)}{P(X > a)} = \frac{1 - F_X(a+h)}{1 - F_X(a)}.$$

Draw a figure to help understand $\{X > a + h\} \cap \{X > a\} = \{X > a + h\}$.

$$\text{(ii) } p_2 = \frac{P(a < X < a+h)}{P(X > a)} = \frac{F_X(a+h) - F_X(a)}{1 - F_X(a)}.$$

$$\text{(iii) } p_3 = \lim_{h \to 0} \frac{[F_X(a+h) - F_X(a)]/h}{1 - F_X(a)} = \frac{F_X'(a)}{1 - F_X(a)} = \frac{f_X(a)}{1 - F_X(a)}$$

Illustration. Let $f(x) = \lambda e^{\lambda x}$, then $F(x) = 1 - e^{-\lambda x}$. Then

$$p_1 = \frac{e^{-\lambda(a+h)}}{e^{-\lambda a}} = e^{-\lambda h};\ p_2 = \frac{[1 - e^{-\lambda(a+h)}] - [1 - e^{-\lambda a}]}{1 - [1 - e^{-\lambda a}]} = 1 - e^{-\lambda h};\ p_3 = \frac{\lambda e^{-\lambda a}}{e^{-\lambda a}} = \lambda.$$

Comments. If $F(x) = (1 - e^{\lambda x})\, u(x - a)$, then $f(x) = F'(x) = (1 - e^{-\lambda x})\, \delta(x - a) + \lambda e^{-\lambda x}(x - c)$.

***Example* 2.** (i) (*Baye's Reversal Rule*). Show that

$$F_X(x \mid A) = [P(A \mid X \le x) \,.\, F_X(x)]/P(A). \quad \text{...(1)}$$

(ii) Show $P(A) = P(A \mid X \le x)\, F(x) + P(A \mid X > x)\, [1 - F(x)]$. ...(2)

Solution. (i) By definition : $F_X(x \mid A) = \dfrac{P\{(X \le x), A\}}{P(A)}$, $P(A \mid X \le x) = \dfrac{P\{X \le x), A\}}{P(X \le x)}$.

Eliminating the number $P\{(X \le x), A\}$ yields (1).

(ii) Let $B = \{X \le x\}$, $\overline{B} = \{X > x\}$, whence $P(B) = F_X(x)$, $P(\overline{B}) = 1 - F_X(x)$.

Substituting these in the Multi-Stage p-Rule : $P(A) = P(A \mid B)\, P(B) + P(A \mid \overline{B})\, P(\overline{B})$ yields (1).

3-54. Probability Density Function

A function f is called a probability density function (p.d.f.) if it satisfies the two conditions :

(1) $f(x) \ge 0$, for all x (Non-negativity)

(2) $\sum_X f(x) = 1$, when X is discrete, or $\int_{-\infty}^{\infty} f(x)\, dx = 1$, when X is continuous. [Normalization]

Note. When X is discrete, f is called probability mass function ; but now-a days the term 'density function' is used both for p.m.f. or p.d.f.

Remarks. In this definition of density function, we have dispensed with introduction of distribution function. We may observe that conditions (1) and (2) are the same as in conditions (ii) of Theorems in Art 3-31 and 3-51.

3-55. Worked-out Problems

Example 1. (a) Can the following be probability functions :

$$f_1(x) = \begin{cases} \frac{3}{4}, & x=1 \\ \frac{1}{4}, & x=2 \\ 0, & \text{elsewhere} \end{cases}; \qquad f_2(x) = \begin{cases} \frac{1}{2}, & x=-1 \\ \frac{2}{3}, & x=1 \\ 0, & \text{elsewhere} \end{cases}$$

$$f_3(x) = \begin{cases} -\frac{1}{3}, & x=2 \\ \frac{1}{3}, & x=3 \\ 0, & \text{elsewhere} \end{cases} \qquad f_4(x) = \begin{cases} 2x, & 0<x\le 1 \\ 4-2x, & 1<x<2 \\ 0, & \text{elsewhere} \end{cases}.$$

Solution. (i) Here $f_1(x) \ge 0 \ \forall\ x$, $\Sigma f_1(x) = 1$. Hence f_1 is a p.m.f. for some variate X.

(ii) Here $f_2(x) \ge 0 \ \forall\ x$, $\Sigma f_2(x) = \frac{1}{2}+\frac{2}{3}=\frac{7}{6} \ne 1$, f_2 is not a p.m.f.

(iii) Here $f_3(2) = -\frac{1}{3}<0$, hence f_3 cannot be a p.d.f. $[\because f(x) \ge 0 \ \ \forall\ x]$

(iv) Here f_4 is not a p.d.f. since $f_4(x) \ge 0 \ \forall\ x$, but

$$\int_{-\infty}^{\infty} f_4(x)\,dx = \int_0^1 2x\,dx + \int_1^2 (4-2x)dx = 1+4-3 \ne 1.$$

Example 2. Prove or disprove : If $f_1(x)$ and $f_2(x)$ are p.d.f's and if $\theta_1 + \theta_2 = 1$, then

$$g(x) = \theta_1 f_1(x) + \theta_2 f_2(x) \text{ is a p.d.f.} \qquad \text{(mixture of densities)}$$

Solution. Define $f_1(x) = 1$, $0 \le x \le 1$; $f_2(x) = 1$, $1 \le x \le 2$. Choose $\theta_2 = 2$, $\theta_1 = -1$, then

$$g(x) = \theta_1 f_1(x) + \theta_2 f_2(x) = 2f_2(x) - f_1(x).$$

Now $$g\left(\tfrac{1}{2}\right) = 2f_2\left(\tfrac{1}{2}\right) - f_1\left(\tfrac{1}{2}\right) = -1. \quad \left\{\left[f_2\left(\tfrac{1}{2}\right) = 0 \text{ as } \tfrac{1}{2} \notin [1, 2]\right\}\right.$$

We conclude that g is not a p.d.f.

Remark. If we impose $0 < \theta_1, \theta_2 < 1$, then g is *always* a p.d.f.

Example 3. For the density $f_X(x) = Ke^{-ax}(1 - e^{-ax})\, I_{(0,\,\infty)}(x)$, find the normalizing constant K, $F_X(x)$ and evaluate $P(X > 1)$.

Solution. (i) For norming constant K,

$$1 = K\int_0^{\infty} e^{-ax}(1-e^{-ax})\,dx = K\left[\int_0^{\infty} e^{-ax}\, x^{1-1}\,dx - \int_0^{\infty} e^{-2ax}\, x^{1-1}\,dx\right] = K\left[\frac{1}{a} - \frac{1}{2a}\right]$$

Thus, $K = 2a$; and so $f(x) = 2a\,(e^{-ax} - e^{-2ax})$.

(ii) $F_X(x) = \int_0^x f(t)\,dt = 2a\int_0^x (e^{-at} - e^{-2at})\,dt = 1 - 2e^{-ax} + e^{-2ax}$, $0 < x < \infty$.

(iii) $P(X > 1) = 2a\int_1^{\infty} (e^{-ax} - e^{-2ax})\,dx = 2[e^{-a} - \frac{1}{2}e^{-2a}] = 2e^{-a} - e^{-2a}$.

Note. $P(X > 1) = 1 - P(X \le 1) = 1 - F(1) = 1 - [1 - 2^{-a} + e^{-2a}] = 2e^{-a} - e^{-2a}$.

Example 4. A variate X has p.d.f. $f(x) = Cxe^{-x^2/8}$, $0 \le x < \infty$. Determine C.

If n independent observations $X_1, X_2, \ldots, X_n$ of X are made, what is the probability that $Y > 4k$, where k is a positive number and $Y = \max.\ \{X_1, X_2, \ldots, X_n\}$.

Solution. We determine C using normality $\int f(x)\,dx = 1$. Thus,

$$1 = C\int_0^\infty xe^{-x^2/8}\,dx = 4C\int_0^\infty e^{-t}\,dt = 4C, \left[t = \frac{x^2}{8},\ x\,dx = 4\,dt\right]$$

Thus $C = \frac{1}{4}$. Now,

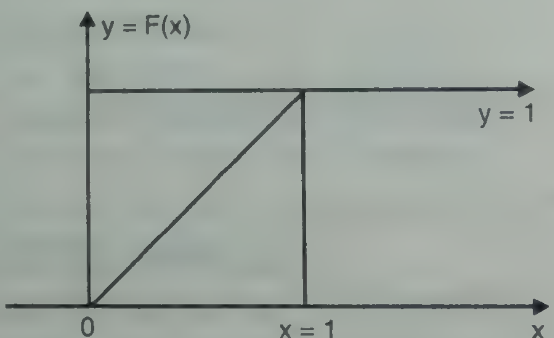
$$\begin{aligned} P\{Y > 4k\} &= 1 - P\{Y \le 4k\} = 1 - P\{\max.(X_1, X_2, \ldots, X_n) \le 4k\} \\ &= 1 - P\{X_1 \le 4k, \ldots, X_n \le 4k\} \\ &= 1 - [P(X_1 \le 4k)]^n \quad [\because X_i \text{ are i.i.d. and Product Rule}] \quad \ldots\text{(i)} \\ P\{X_1 \le 4k\} &= \int_0^{4k} \frac{1}{4} xe^{-x^2/8}\,dx = \int_0^{2k^2} e^{-t}\,dt = (1 - e^{-2k^2}). \end{aligned}$$

Making substitutions into (i) we get $P(Y > 4k) = 1 - (1 - e^{-2k^2})^n$.

Problems with Solutions Provided at the End of the Text

1*. Let F be the c.d.f. of a continuous variate X defined by

$$F(x) = \begin{cases} 0, & x \le 0 \\ x, & 0 \le x \le 1 \\ 1, & x > 1 \end{cases}.$$



Find the p.d.f. of X. Use both F and f to evaluate

$P\left(\frac{1}{3} \le X \le \frac{1}{2}\right)$ and $P\left(\frac{1}{2} \le X \le 2\right)$.

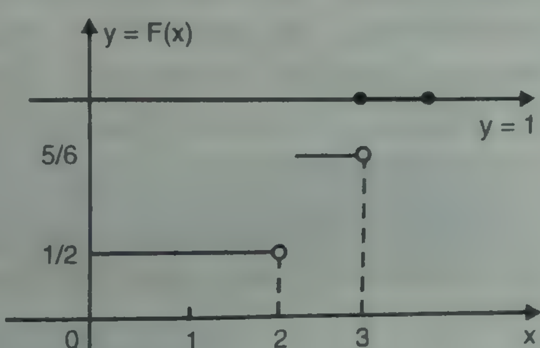
2*. The c.d.f. 'F' of a continuous variate X is given by

$$F(x) = \begin{cases} 0, & x < 0 \\ x^2, & 0 \le x \le \frac{1}{2} \\ 1 - 3(3-x)^2/25 & \frac{1}{2} \le x < 3 \\ 1, & x \ge 3. \end{cases}$$

Find the p.d.f. of X and evaluate $P(|X| \le 1)$ and $P\left(\frac{1}{3} \le X < 4\right)$, using both F and f.

3*. Obtain the density function for the variate X whose distribution is specified by

$$F_X(x) = \begin{cases} 0, & x < 0 \\ \frac{1}{2}, & 0 \le x < 2 \\ \frac{5}{6}, & 2 \le x < 3 \\ 1, & x \ge 3 \end{cases}$$



4*. The d.f. 'F' of a continuous variate X is defined by

$$F(a) = 0.5,\ F(b) = 0.7,\ F(c) = 0.8.$$

Evaluate $P(a \le X \le b)$, $P(b \le X \le c)$. Calculate also the probability that two independent observations X_1 and X_2 will lie between $-\infty$ to a and c to ∞.

5*. Let X be continuous with p.d.f.

$$f(x) = k(x - x^2),\ a < x < b,\ k > 0. \quad \ldots(1)$$

Find the possible values of a, b and k.

6*. Determine the distribution function $F(x)$, for the following density function and calculate $F(3)$:

$$f(x) = \begin{cases} x/3, & 0 < x \le 1 \\ 5(4-x)/27, & 1 < x \le 4 \\ 0, & \text{elsewhere} \end{cases}$$

7*. X is a continuous variate with frequency : $f(x) = \frac{1}{2} k e^{-k|x|}, k > 0, (-\infty, \infty)$. Verify that f is a density function and derive $F(x)$.

8*. Life X in hours of a certain radio-tube has p.d.f.

$$f(x) = 100/x^2,\ x \ge 100\ ; \quad f(x) = 0,\ x < 100.$$

Find the distribution function of X. What is the probability that none of three such tubes in a given radio-set will have to be replaced during the first 150 hours of operation ? Find the probability that all three of the original tubes will have been replaced during the first 150 hours.

9*. A wood-boring larvae travels a distance X(cm) along a long rod as per p.d.f.

$$f(x) = \theta/1(+x)^{\theta+1},\quad x \ge 0,\ \theta > 0\ ;\ f(x) = 0, \text{ otherwise}$$

where θ is the measure of hardness (m.h.) of wood.

(a) Check that θ is a sensible m.h. of wood by showing that $P(X \ge a)$ decreases as θ increases.

(b) What m.h. of wood ensures a probability of more than 99% that the larva is within 1 cm of the starting point ?

(c) Show that the m.h. of wood which maximizes the probability that the larva has travelled between 1 and 2 cm from the starting point is [by ($\ln 3/\ln 2$)]/$\ln$ (3/2).

3-60. *Some Standard Continuous Distributions*

Some commonly occuring probability density function [p.d.f.] have acquired definite terminology. We quote these densities which will form a part of solid core in the ongoing study of Probability Theory. Notation follows the name of distribution.

1. ***Uniform (or rectangular) distribution.*** $X \sim U(a, b)$ or $X \sim \text{unif}(a, b)$.

$f(x) = 1/(b-a),\ a \le x \le b,\ f(x) = 0$, elsewhere, $[b > a, -\infty < a < \infty]$.

$F_X(x) = 0\ \ I(x < a) + [(x-a)/(b-a)]\ I(a \le x < b) + 1\ I(x \ge b)$.

2. ***Normal (or Gaussian) distribution.*** $X \sim N(\mu, \sigma^2)$.

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}, \quad \sigma > 0, |\mu| < \infty, -\infty < x < \infty.$$

$$F_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-(t-\mu)^2/2\sigma^2} dt.$$

The *Standard* or *unit normal distribution*, notated : $X \sim N(0, 1)$

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, -\infty < x < \infty,\ F(x) \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2}\, dt = \Phi(x)$$ [Standard notation]

3. ***Log-normal distribution.*** $X \sim \ell n\ N(\mu, \sigma^2)$

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \frac{1}{x} . \exp\left[-\frac{(\ell n\, x-\mu)^2}{2\sigma^2}\right],\ x > 0 .$$

Terminology. $X \sim \ell n$ (Named distribution) iff $\ell n\ X \sim$ (Named distribution).

4. ***Gamma distribution.*** $X \sim$ gam (α, λ)

$$f(x) = \frac{e^{-\lambda x}\, x^{\alpha-1}}{\Gamma(\alpha)/(\lambda)^{\alpha}} . x > 0,\quad f(x) = 0, \text{if } x \le 0$$

5. ***Exponential distribution.*** $X \sim$ Expo (λ).

$f(x) = \lambda e^{-\lambda x}$, $x > 0$, $f(x) = 0$, elsewhere, [Expo $(\lambda) \equiv$ gam $(1, \lambda)$]

6. ***Chi-square distribution.*** $X \sim \chi^2_{(n)}$. [n is called degrees of freedom]

$$f(x) = \frac{e^{-x/2} . x^{(n/2)-1}}{\Gamma(n/2) . 2^n}, x > 0,\ f(x) = 0, \text{ elsewhere}$$ [$\chi^2_{(n)} \equiv$ gam $(n/2, 1/2)$]

7. ***Beta distribution of the first kind.*** $X \sim B_1\ (a, b)$.

$$f(x) = x^{a-1}(1-x)^{b-1} / B(a,b),\ a > 0, b > 0, 0 < x < 1.$$

8. ***Beta distribution of the second kind.*** $X \sim B_{II}(a, b)$.

$$f(x) = \frac{1}{B(a,b)} \frac{x^{a-1}}{(1+x)^{a+b}},\ a > 0,\ b > 0,\ 0 < x < \infty.$$

9. ***Laplace (or double-exponential distribution.*** $X \sim$ Lap (a, λ)

$$f(x) = \tfrac{1}{2} \lambda\, e^{-\lambda |x-a|}, -\infty < x < \infty,\ \lambda > 0,\ |a| < \infty.$$

10. ***Cauchy distribution.*** $X \sim$ Chy (a, b).

$$f(x) = \frac{b}{\pi} . \frac{1}{b^2 + (x-a)^2} .\ b > 0,\ |a| < \infty,\ -\infty < x < \infty.$$

11. ***Student's t-distribution.*** $X \sim t(n)$.

$$f(x) = \frac{1}{B(\frac{1}{2}, \frac{n}{2})} \frac{1}{\sqrt{n}} \left(1 + \frac{t^2}{n}\right)^{-(n+1)/2},\ -\infty < t < \infty.$$

12. ***Snedecor-Fisher F-distribution.*** $X \sim F_{m\ .\ n}$ [or $X \sim F(m, n)$]

$$f(x) = \frac{(m/n)^{m/2}}{B(\frac{1}{2} m, \frac{1}{2} n)} . \frac{x^{(m/2)-1}}{[1 + (mx/n)]^{(m+1)/2}}, m > 0,\ n > 0;\ 0 < x < \infty.$$

13. ***Rayleigh distribution.*** $X \sim$ Ryl (a, λ)

$$f(x) = 2\lambda(x-a)\, e^{-\lambda(x-a)^2}, x \ge a,\ f(x) = 0,\ x < a.$$

$$F(x) = [1 - e^{-\lambda(x-a)^2}]\, I(x \ge a)^2 + 0\, I(x, a).$$

14. ***Logistic distribution.*** $f(x) = \lambda \exp[-\lambda(x-a)]/[1+\exp-\lambda(x-a)]^2$; $\lambda > 0$, $|a| < \infty$, $|x| < \infty$.

15. ***Burr distribution.*** $f(x) = ab\, x^{a-1}/(1+x^a)^{b+1}$; $a > 0$, $b > 0$, $x \geq 0$.

16. ***Pareto distribution.*** $f(x) = \theta k^{\theta}/x^{(\theta+1)}$; $\theta > 0$, $x \geq k > 0$.

17. ***Weibull distribution.*** $f(x) = \lambda a\, x^{a-1} \exp(-\lambda x^a)$, $\lambda > 0$, $a > 0$, $c > 0$.

18. ***Gumbel or extreme-value distribution.***

$$f(x) = \lambda \exp\{-\lambda(x-a) - e^{-\lambda(x-a)}\}\ ;\ F(x) = \exp\{-e^{-\lambda(x-a)}\},\ \lambda > 0,\ |a| < \infty,\ |x| < \infty.$$

19. ***Pearsonian System of distributions.***

$$\frac{dy}{dx} - \frac{L}{Q}y = 0, \quad \text{i.e.} \quad \frac{dy}{dx} = \frac{(x+b)\,y}{a_0 + a_1 x + a_2 x^2}. \quad [L = \text{linear},\ Q = \text{quadratic}]$$

The four constants a_0, a_1, a_2, b can be expressed in terms of the first four moments of the density function which satisfies this differential equation.

Exercise 3(c)

1. Which of the following are genuine (discrete/continuous) density functions ?

(a) $f(x) = \begin{cases} 1/12, & x = 0 \\ 7/12, & x = 1 \\ 1/4, & x = 2 \\ 0, & \text{elsewhere} \end{cases}$ (b) $f(x) = \begin{cases} 2^{-|x|}, \\ x = \ldots -1, 0, 1, 2, \ldots \\ 0, \quad \text{otherwise} \end{cases}$

(c) $f(x) = \begin{cases} \sin x, & -\pi/2 < x < \pi \\ 0, & \text{otherwise} \end{cases}$ (d) $f(x) = \begin{cases} \frac{1}{2}|\sin x|, & -\frac{1}{2}\pi < x < \frac{1}{2} \\ 0, & \text{otherwise} \end{cases}$

(e) $f(x) = \begin{cases} e^{-|x|}\pi^{-1}, & -3 < x < 1, \\ 0, & \text{elsewhere} \end{cases}$ (f) $f(x) = \begin{cases} |x|, & -1 < x < 1 \\ 0, & \text{otherwise} \end{cases}$

(g) $f(x) = \begin{cases} \dfrac{\ln x}{\int_{1/2}^{3} \ln x\, dx}, & \frac{1}{2} < x < 3 \\ 0, & \text{otherwise} \end{cases}$ (h) $f(x) = \begin{cases} \frac{2}{3}\left(\frac{2}{3}\right)^x, & x = 0, 1, 2, \ldots \\ 0, & \text{elsewhere} \end{cases}$

[**Ans.** No, No, No, Yes, No, Yes, Yes, Yes]

2. Find the value of the constant k in the following cases so that $f(x)$ represents a valid density function :

(a) $f(x) = kx^2$, $0 \leq x \leq a$, $a > 0$ (b) $f(x) = kx^2$, $-k \leq x \leq k$

(c) $f(x) = ke^{-\lambda x}$, $x \geq 0$, $\lambda > 0$ (d) $f(x) = kxe^{-\lambda x}$, $x \geq 0$, $\lambda > 0$

(e) $f(x) = k/(1+x)^2$, $0 < a < x < b$

(f) $f(x) = \begin{cases} 2k/(1+x)^2, & 0 < x < 1 \\ kx^2, & 2 < x < 3 \\ 0, & \text{elsewhere} \end{cases}$ (g) $f(x) = \begin{cases} kx, & 0 \leq x \leq 1 \\ k, & 1 \leq x \leq 2 \\ k(3-x), & 2 \leq x \leq 3 \end{cases}$

[**Ans.** $3/a^3$, $\left(\frac{3}{2}\right)^{\frac{1}{4}}$, λ, λ^2, $(1+a)(1+b)/(b-a)$; $\frac{6}{41}$]

3. A continuous variate X follows the probability law

$f(x) = kx^2$, $0 \leq x \leq 1$; $f(x) = 0$, elsewhere.

Determine k and show that

$$P\{0.2 \le X \le 0.5\} = 0.117 \,;\, P(X \le 0.3) = 0.027 \,;\, P(X > \tfrac{3}{4} \mid x > \tfrac{1}{2}) = 37/56.$$

4. Let $f(x) = \frac{1}{4}, |x| < 2 \,;\, f(x) = 0$, elsewhere, Prove that

$P(X < 1) = \frac{3}{4}$, $P(|X| > 1) = 1/2$, $P(2X + 3 > 5) = 1/4$.

5. Show that the following are genuine p.d.f's .

(i) $f(x) = e^{-x}, x \ge 0$ (ii) $g(x) = 2e^{-2x}, x \ge 0$

(iii) $h(x) = (1 + \theta) f(x) - \theta g(x) \,;\, x \ge 0, 0 < \theta < 1.$

6. Mixtures of probability masses. If f and g are any two p.d.f.'s, show that $af + bg$ is also a p.d.f. where a and b are non-negative constants with $a + b = 1$.

7. A variate X has the following p.d.f. : $f(x) = kx(2 - x), 0 \le x \le 2$.

(a) Find k. Show that the frequency curve is symmetrical and find the axis of symmetry.

(b) Find the equation to the cumulative distribution and sketch its form.

8. Verify that $F_X(x)$ defined as under are c.d.f.s and obtain their correspondin p.d.f.'s

(a) $F(x) = \begin{cases} 0, x < -a \\ \frac{1}{2}[1 + (x/a)], |x| \le a \\ 1, \quad x > a \end{cases}$ (b) $F(x) = \begin{cases} 0, & x < -2 \\ \frac{1}{2} & -2 \le x < a \\ 1, & x \ge 0 \end{cases}$

Show : $P(-a/2 < X \le a/2) = \frac{1}{2}$. Show $P(|X| \le = \frac{1}{2})$.

(c) $F(x) = \begin{cases} 0, & x < 3 \\ \frac{1}{3}, & 3 \le x < 4 \\ \frac{1}{2}, & 4 \le x < 5 \\ \frac{2}{3}, & 5 \le x < 6 \\ 1, & x \ge 6 \end{cases}$ (d) $F(x) = \begin{cases} 0, & x < 0 \\ \frac{1}{2}x, & 0 \le x < 1 \\ 1 - (\frac{1}{2}x), & x \ge 1 \end{cases}$

Use $f(x)$ to show $P(3 < X \le 5) = \frac{1}{3}$

(e) $F(x) = \begin{cases} 0, & x < 0 \\ x^2, & 0 \le x < \frac{1}{2} \\ 1 - 3(1 - x)^2, & \frac{1}{2} \le x < 1 \\ 1, & x \ge 1 \end{cases}$ (f) $F(x) = \begin{cases} 0, & x < 1 \\ 1 - \frac{1}{4}e^{-x}, & x \ge 0 \end{cases}$

Show $P(X = 0) = \frac{3}{4}$ and $P(X \ge 0) = 1$.

(g) $F(x) = \begin{cases} 0, & x < 1 \\ x - x \ln x, & 0 \le x < 1 \\ 1, & x \ge 1 \end{cases}$ (h) $F(x) = \begin{cases} 1 - (1 + x)e^{-x}, & x \ge 0 \\ 0, & x < 0 \end{cases}$

(i) $F(x) = \dfrac{1}{2} + \dfrac{x}{2(1 + |x|)}, -\infty < x < \infty$. Find points where $F'(x) = f(x) = 0$

[**Ans.** (a) $U[-a, a]$, (b) $f(0) = f(-2) = \frac{1}{2}$, (c) $f(3) = f(6) = \frac{1}{3}, f(4) = f(5) = \frac{1}{6}$.

(d) $f(x) = \frac{1}{2} I_{(0,1)} + (2x^2)^{-1} I_{(x \ge 1)}$ (e) $f(x) = 2x I_{(0,\frac{1}{2})} + 6(1 - x) I_{(\frac{1}{2},1)}$

(f) $f(x) = \frac{3}{4}, x = 0, f(x) = \frac{1}{4}e^{-x}, x > 0$

(g) $f(x) = -\ln x, 0 < x < 1$, (h) $f(x) = xe^{-x}, x > 0$].

9. (a) The logistic distribution is defined by the c.d.f.

$F(x) = [1 + e^{-(ax+b)}]^{-1}$; $-\infty < x < \infty, a > 0, -\infty < b < \infty$

Show that : $f(x) = aF(x)[1 - F(x)]$.

(b) Show that $F(x) = (1 + e^{-x})^{-1}, |x| < \infty$, is a c.d.f. of a r.v. X. Find the p.d.f. of X and show that it is symmetric about $x = 0$.

10. Find the distribution function $F_X(x)$ for the density function :

(a) $f(x) = \begin{cases} \frac{1}{3}, & x = 0, 1, 2 \\ 0, & \text{otherwise} \end{cases}$ (b) $f(x) = \begin{cases} \frac{1}{3}, & x = -3 \\ \frac{1}{6}, & x = 0 \\ \frac{1}{3}, & x = 4 \end{cases}$

(c) $f(x) = \begin{cases} 1/(b-a), & a < x < b \\ 0, & \text{otherwise} \end{cases}$ (d) $f(x) = \begin{cases} \frac{3}{4}(1 - x^2), & |x| < 1 \\ 0, & \text{otherwise} \end{cases}$

Show : $P(-\frac{1}{4} \le X \le \frac{1}{2}) = \frac{135}{256}$.

(e) $f(x) = \begin{cases} \frac{1}{2}, & -1 \le x \le 0 \\ \frac{1}{4}(2 - x), & 0 < x < 2 \\ 0, & \text{elsewhere} \end{cases}$ (f) $f(x) = \begin{cases} x, & 0 < x \le 1 \\ 2 - x, & 0 \le x < 2 \\ 0, & \text{elsewhere} \end{cases}$

Show : $P\left(-\frac{1}{2} < x \le 1\right) = \frac{5}{8}$; $P(X > 1) = \frac{1}{8}$; $P\left(X > -\frac{1}{2} < |X \le 1\right) = \frac{5}{7}$.

(g) $f(x) = \begin{cases} 0, & x \le -a \\ (a + x)/a^2, & -a < x \le 0 \\ (a - x)/a^2, & 0 < x \le a \\ 0, & x \ge a \end{cases}$

[**Ans.** (a) $F(x) = 0\, I(x < 0) + (\frac{1}{3})\, I_{[0, 1)} + (\frac{2}{3})\, I_{[1, 2)} + 1 I_{[2\, \infty)}$.

(b) $F(x) = 0 I(x < -3) + \frac{1}{2} I(-3 \le x < 0) + \frac{2}{3} I(0 \le x < 4) + 1\, I(x \ge 4)$.

(c) $F(x) = 0\, I(x < a) + [(x - a)/(b - a)]\, I(a \le x < b) + 1 .\, I(x \ge b)$.

(d) $F(x) = I(x < -1) + \frac{1}{4}(x + 1)(2 + x - x^2)\, I(-1 \le x < 1) + 1\, I(x \ge 1)$.

(e) $F(x) = 0\, I(x < -1) + \frac{1}{2}(1 + x)\, I(-1 \le x < 0) + \frac{1}{2} + \frac{1}{2} x(1 - \frac{1}{4}x)\, I(0 \le x < 2) + 1\, I(x \ge 2)$.

(f) $F(x) = 0\, I(x < 0) + \frac{1}{2} x^2\, I(0 \le x < 1) + [\frac{1}{2} + \frac{1}{2}(x - 1)(3 - x)]\, I\, (1 \le x < 2) + 1 I(x \ge 2)$.

(g) $F(x) = 0\, I(x < -a) + \frac{1}{2}(a + x)^2\, I(-a < x < 0) + \frac{1}{2} a^{-2} (a - x)^2\, I(0 < x < a) + 1\, I(x \ge a)$.

11. Show that $f(x)$ specified as under is not a p.d.f. for any value of the constant a : $f(x) = 3ax^2 - 4x - 5$, $2 \le x \le 3$; $f(x) = 0$, elsewhere.

12. The mileage X in thousands of miles which car owners get with certain kind of tyre is a variate with p.d.f.

$$f(x) = (1/20)\, e^{-x/20}, x > 0; f(x) = 0, x \le 0.$$

Find the probabilities that one of these tyres will last :

(a) At least 30,000 miles. (b) At most 10,000 miles.

(c) Anywhere from 16000 to 24000 miles.

13. A country filling station is supplied with gasoline once a week. Its weekly volume X of sales in thousands of gallons is distributed by $f(x) = 5(1-x)^4$, $0 < x < 1$. Prove that the capacity of its tank, in order that the probability that its supply will be exhausted in a given week shall be 0.01, is $[1 - (0.01)^{1/5}]$ thousand of gallons.

14. A bombing plane carrying three bombs flies directly above a rail-road track. If a bomb falls within 40 feet of track, the track will be sufficiently damaged to disrupt the traffic. Within a certain bomb site the points of impact of a bomb have the p.d.f.

$$f(x) = \begin{cases} (100+x)/10{,}000, & -100 \le x \le 0 \\ (100-x)/10{,}000, & 0 \le x \le 100 \\ 0, & \text{elsewhere} \end{cases}$$

where x represents the vertical deviation (in feet) from the aiming point, which is the track in this case. Find the distribution function. If all the bombs are used what is the chance that the track will be damaged ?

15. For a certain machine, the time X from restart after a service to next breakdown, is a variate with p.d.f. $f(x) = 2/(1+x)^3, x \ge 0$; $f(x) = 0$, elsewhere.

Such a machine is routinely inspected at time t after a restart and is found to have broken down already. What is the chance that it has been out of action for at least time kt $(0 < k < 1)$?
Another machine of the same type is found to be operating at time t after restart due to have broken down by the end of a further time t. What is the chance that it has been out of action for at least time kt $(0 < k < 1)$? $\left[\textbf{Ans. } p_1 = \dfrac{1-[1+(1-k)t]^{-2}}{1-(1+t)^{-2}}, \; p_2 \; \dfrac{(1+t)^{-2}-[1+(2-k)t]^{-2}}{(1+t)^{-2}-(1+t)^{-2}}\right]$

16. Time (in minutes) that a person has to wait at a certain station for a train is found to have the c.d.f. : $F(x) = 0$, or $x \le 0$; $F(x) = x/2, 0 \le x < 1$.

$F(x) = \frac{1}{2}, \quad 1 \le x < 2; F(x) = x/4, \quad 2 \le x < 4, \quad F(x) = 1, x \ge 4.$

(a) Is the c.d.f. continuous ? If so, find the formula for its p.d.f.

(b) Show that the chance that a person will have to wait (i) more than 3 minutes is $\frac{1}{4}$, (ii) less than 3 minutes is $\frac{3}{4}$, (iii) between 1 and 3 minutes is $\frac{1}{4}$.

(c) Show that the conditional probability that a person will have to wait for a train for

(i) more than 3 minutes given that it is more than 1 minutes is $\frac{1}{2}$.

(ii) less than 3 minutes, given that it is more than 1 minute is $\frac{1}{2}$.

17. The length of life X (in hours) of a certain type of light bulb may be supposed to be a continuous variate with p.d.f. $f(x) = a/x^3$, $1500 < x < 2500$; $f(x) = 0$ elsewhere.

Determine the constant a and the distribution function X. Also find

$$P\{1800 < X < 2000\}, P(1700 < X < 1900) \mid 1500 < X < 2000)\}.$$

18. A batch of small calibre ammunition is accepted as satisfactory if none of a sample of 5 shots falls more than 2 feet from the centre of the target at a given range. The distance from the centre of the target to a given impact point, denoted by variate R, has the p.d.f., $f(r) = 2kr\,e^{-r^2}$, $0 \le r \le 3$. Find the constant k and show that the probability that a given batch will be acceptedis $[(1-e^{-4})/(1-e^{-9})]^5$.

19. (a) If $x \geq 20$, the probability that a man aged x marries at time t from the present time is $15\, dt/(t+x)^2$. Show that P {person aged 20 eventually marries} $= \frac{3}{4}$.

(b) The probability that a person will die in the time interval (t_1, t_2) is given by

$$P(t_1 \leq t \leq t_2) = \int_{t_1}^{t_2} f(t)\, dt.$$

The function $f(t)$ is determined from long records and is given by

$$f(t) = kt^2 (100 - t)^2, 0 \leq t \leq 100 \; ; \; f(t) = 0, \text{ elsewhere.}$$

Show that the probability that a person will die between the ages of 60 and 70 assuming that he lived to be 60 is 0.486.

20. Let X be continuous with c.d.f. F and p.d.f. f. Show that $Y = X^2$ is also continuous and relate f_Y and F_Y to F and f.

21. A continuous r.v. X has a differentiable c.d.f. $F(x)$ satisfying :

(a) $F(x) = 0$, for $x \leq 0$. (b) $0 < F(x) < 1$, for $0 < x < 1$. (c) $F(x) = 1$, for $x \geq 1$.

Show that the polynomial p.d.f. $f(x)$ for X, which has lowest degree is

$$f(x) = 6x(1-x), \; 0 \leq x \leq 1 \; ; f(x) = 0, \text{ elsewhere.}$$

3-70. *Determination of the p.m.f. or p.d.f. of a Function of a Single r.v.*

(i) Let X and Y be discrete r.vs., such that $Y = g(X)$. Then p.m.f. of Y is given by

$$f_2(y) = \Sigma_A f_1(x), \; A = \{x : g(x) = y\}.$$

Proof. $f_2(y) = P\{Y = y\} = P\{g(X) = y\} = \Sigma_A P(X = x) = \Sigma_A f_1(x)$

Illustration. Find p.m.f. of $Y = g(X)$, where

(a) $Y = -X$, (b) $Y =$ sign X, (c) $Y = \max\{0, X\}$, (d) $Y = \max\{0, -X\}$, (e) $Y = |X|$.

Solution. (a) $f_Y(x) = f_X(-x)$ $\quad [Y = -x \Rightarrow x = -y]$

(b) $f_Y(x) = \Sigma_A f_1(x) \; ; \; x = 1, A = \{x > 0\}$ $\quad Y = \text{sign } X = X/|X|, X \neq 0$

$\quad f_1(0) \; ; \; x = 0$ $\quad Y = 0, X = 0$

$\quad \Sigma_B f_1(x), \; x = -1, B = \{x < 0),$

(c) $f_2(x) = f_1(x), \quad x > 0$

$\quad \Sigma_A f_1(x) \; x = 0, A = \{x \leq 0\}$ $\quad \max(0, X) = X^+$

(d) $f_2(x) = f_1(x) \; ; x > 0$

$\quad \Sigma_A f_1(x); \; x = 0$ $\quad \max\{0, -X\} = X^-$

(e) $f_2(x) = f_1(x) + f_1(-x), \; x \neq 0$ $\quad |X| = X^+ + X^-$

$\quad f_1(x), \quad x = 0$

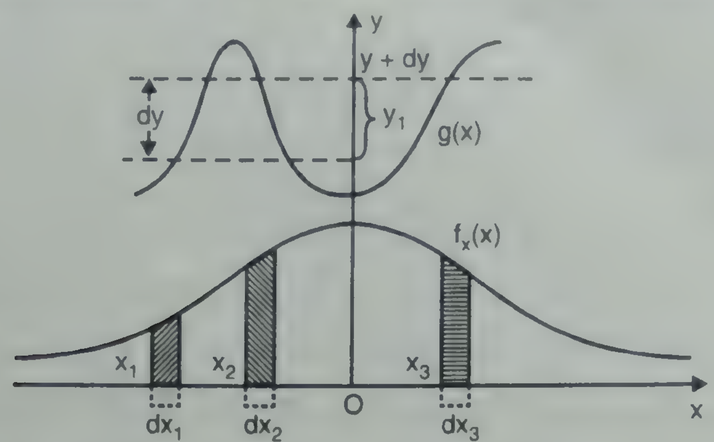
(ii) Let $Y = g(X)$, so that we need the density $f_Y(y)$ of Y in terms of the density $f_X(x)$ of X. We assume that X is of continuous type and $g(x)$ is continuous and it does not equal a constant over any interval. This means that, for a given y, the equation $y = g(x)$ has at most a countable number of roots $x_1, x_2, x_3, \ldots$

3-71. *Fundamental Formula*

If $x_1, x_2, \ldots, x_n, \ldots$ are all *real* roots of $y = g(x)$, $[y = g(x_1) = g(x_2) = \ldots]$, then the density function of Y is

$$f_Y(y) = \frac{f_X(x_1)}{|g'(x_1)|} + \frac{f_X(x_2)}{|g'(x_2)|} + \ldots + \frac{f_X(x_n)}{|g'(x_n)|} + \ldots [g'(x) = dg(x)/dx]. \quad \ldots(1)$$

If for a certain y, the equation $y = g(x)$ has no real root, then $f_Y(y) = 0$.



Proof. To simplify the matter, assume that for the given y, the equation $y = g(x)$ has three roots x_1, x_2, x_3. Then

$$f_Y(y)\, dy = P\{y < Y < y + dy\},\ dy > 0.$$

Thus to find $f_Y(y)$, it suffices to find all values of x such that $y < g(x) < y + dy$. And this is true for (see Fig.) the following :

$A_1 = \{x_1 < X < x_1 + dx_1\}$, $A_2 = \{x_2 + dx_2 < X < x_2\}$, $A_3 = \{x_3 < X < x_3 + dx_3\}$ where $dx_i > 0$, $i = 1, 2, 3$. Hence

$P\{y < Y < y + dy\} = P(A_1 \uplus A_2 \uplus A_3) = P(A_1) + P(A_2) + P(A_3)$, [Shaded area shown]

Now : $P(A_1) = P(x_1 < X < x_1 + dx_1) = f_X(x_1)\, dx_1$. $P(A_2) = f_X(x_2)\, |\, dx_2\, |$, $P(A_3) = f_X(x_3)\, dx_3$.

And $\quad g'(x_i)\, dx_i = dy,\ i = 1, 2, 3$. Thus,

$$f_Y(y)\, dy = \frac{f_X(x_1)}{|g'(x_1)|} dy + \frac{f_X(x_2)}{|g'(x_2)|} dy + \frac{f_X(x_3)}{|g'(x_3)|} dy.$$

The result (1) now follows from the extension of this case.

3-72. *Applications of Fundamental Formula*

(1) Let $Y = g(X) = aX + b$. (Linear From)

Here $y = ax + b \Rightarrow x = (y - b)/a$, for all y (single solution), $g'(x) = a$. Hence

$$f_Y(y) = (|\, a\, |) - 1\, f_X\,[y - b/a].$$

Illustration. If X is uniformly distributed over $[x_1, x_2]$, then

$f_X(x) = 1/(x_2 - x_1),\ x_1 < x < x_2, \qquad f_x(x) = 0$, otherwise $\qquad [x = (y - b)/a]$

Hence, Y is $U(ax_1 + b,\ ax_2 + b)$.

(2) $\qquad Y = g(X) = a/X$, (Hyperbolic form)

Here $y = a/x$, i.e. $x = a/y$, $\forall\ y$ (single solution). Since $|\, g'(x)\, | = a/x^2\ |\ y^2/|\ a\ |$, the density is $f_Y(y) = \dfrac{|a|}{|y_2|} f_Y\left(\dfrac{a}{y}\right)$.

Illustration. If X has a Cauchy distribution, defined by

$$f(x) = \frac{b}{\pi(x^2 + b^2)}, (-\infty, \infty) \qquad (b > 0 : \text{parameter})$$

$$Y = a/X \Rightarrow f_Y(y) = \frac{|a|}{y^2}\cdot\frac{b}{\pi}\cdot\frac{1}{b^2+(a/y)^2} = \frac{|a|/b\pi}{y^2+(a/b)^2}.$$ Thus $Y \sim \text{Chy}\ (0, |a|/b)$.

(3) Let $Y = g(X) = aX^2$ (Square Form : Parabola)

Here $y = ax^2$ so $x = \pm\ (y/a)^{1/2}$ $\forall\ y > 0$ (two solutions)

$$g'(x) = 2a\,x, \text{ so that } g'(x_1) = 2ax_1 = 2\sqrt{ay},\ g'(x_2) = -2\sqrt{ay}.$$

$\therefore\ f_Y(y) = (2\sqrt{ay})^{-1}\{f_X\,[(y/a)]^{1/2} + f_X[-(y/a)^{1/2}]\},\ y>0$. If $y < 0$, then, $f_Y(y) = 0$.

The distribution function $F_Y(y)$ is given by

$$F_Y(y) = P\{Y \le y\} = P\{-\sqrt{y/a} < X < \sqrt{y/a}\} = F_X(\sqrt{y/a}) - F_X(-\sqrt{y/a}).$$

Exercise. Find the density of $Y = X^2$, when X is (a) $N(\mu, \sigma^2)$, (b) Lap $(\lambda\ ;\ a)$ and (c) Chy (a, b).

(4) $Y = g(X) = a \tan x,\ a > 0$.

Here $y = a \tan x \Rightarrow x_n = \tan^{-1}\ (y/a),\ n = \ldots, -1, 0, 1, \ldots$

Obviously, $y = a \tan x$ has infinitely many solutions for an y. Now

$$g'(x) = a \sec^2 x = a(1 + \tan^2 x) = (a^2 + y^2)/a$$

$$\therefore \quad f_Y(y) = \frac{a}{a^2+y^2}\sum_{n=-\infty}^{\infty} f_X(x).$$

Illustration. Let X have the p.d.f. $f_X(x) = \frac{1}{2}\pi$. Now only two terms in the above equation are different from zero. Hence

$$f_Y(y) = \frac{a}{a^2+y^2}\left(\frac{1}{2\pi}+\frac{1}{2\pi}\right) = \frac{a/\pi}{a^2+y^2}. \qquad [Y \sim \text{Chy}\ (0, a)]$$

3-73. *Worked-out Problems*

***Example* 1.** Let $f(x) = 2x,\ 0 < x < 1$ and $f(x) = 0$, elsewhere, be the p.d.f. of X. Find the distribution function and p.d.f. of $Y = \sqrt{X}$.

Solution. $$P\{Y \le y\} = P\{\sqrt{X} \le y\} = P\{X \le y^2\}$$

$$\therefore \quad F_Y(y) = \int_0^{y^2} 2x\,dx = y^4,\quad 0 \le y \le 1.$$

This gives the distribution function of Y. The p.d.f. of Y is $f_Y(y) = F'_Y(y) = 4y^3,\quad 0 \le y \le 1$.

***Example* 2.** The radius X of a particle is a variate with p.d.f. $f(x) = 3x^2,\ 0 < x < 1$. Find the distribution of the volume of the particle.

Solution. The volume of the particle is $V = (4/3)\ \pi X^3$, so we get

$v = (4/3)\ \pi x^3,\ dv = 4\pi x^2 dx$. Also $0 < x < 1 \quad \Rightarrow \quad 0 < v < (4/3)\ \pi$.

The p.d.f. of V is given by $$g(v) = f(x)\left|\frac{dx}{dv}\right| = 3x^2\cdot\frac{1}{4\pi x^2} = \frac{3}{4\pi},\quad 0 < v < \frac{4}{3}\pi.$$

$\therefore\ g(v) = (3/4\pi), \qquad 0 < v < 4\pi/3\ ;\ g(v) = 0$, elsewhere.

The volume is uniformly distributed.

Problems with Solutions Provided at the End of the Text

1*. Let $Y = |X|$ and $Z = (Y)^{1/2}$. Find the c.d.f. of Y and Z in terms of F_X.

2*. A variate X has p.d.f. $f(x) = 1/x^2$, $x \geq 1$; $f(x) = 0$, $x < 1$. Find the p.d.f. of e^{-X}.

3*. Variate X has the p.d.f. $f(x) = 1/\pi\,(1 + x^2)$, $-\infty < x < \infty$. Find the p.d.f. of $\tan^{-1} X$.

Exercise 3(d)

1. If the c.d.f. and p.d.f. [i.e. $F_X(x)$ and $f_X(x)$] are known, find the distribution and density functions for

   (a) $|X|$, (b) $\frac{1}{2}[X + |X|]$, (c) $a \sin (X + \alpha)$, $a > 0$, (d) e^x, (e) $ae^{-\lambda x}\, u(x)$ where $a, \lambda > 0$;

   [$u(x) = 1$, if $x \geq 0$; $u(x) = 0$, if $x < 0$ unit step function].

2. If X is a random variable, let $Y = h(X)$ when h is defined by

   (i) $h(x) = \begin{cases} b, & x \geq b > 0 \\ x, & |x| < b \\ -b, & x \leq -b \end{cases}$ (ii) $h(x) = \begin{cases} x + b, & x \geq 0 \\ x - b, & x < 0 \end{cases}$

   Show that (i) $F_Y(y) = 0\, I(y < b) + F_X(y)\, I(-b \leq y < b) + 1\, I(y \geq b)$

   (ii) $F_Y(y) = F_X(y + b)\, I(y < -b) + F_X(0)\ I(-b \leq y < b + F_X(y - b)\ I(y \geq b)$, $b > 0$.

3. The p.d.f. of X is given by : $f(x) = \frac{1}{3}$, $-1 < x < 2$; $f(x) = 0$, otherwise. Find the c.d.f and p.d.f. of $Y = X^2$.

   [**Ans.** $F_Y(y) = 0.\ I(y < 0) + \frac{2}{3}\sqrt{y}\, I(0 \leq y < 1) + \frac{1}{3}(1 + \sqrt{y})\, I(1 \leq y < 4) + 1\ I(y \geq 4)$

   $f_Y(y) = \frac{1}{3} y^{-1/2}\ I(0 < y < 1) + \frac{1}{6} y^{-1/2}\ I(1 < y < 4)$]

4. Let the p.d.f. of X be $f(x) = \frac{1}{6}, -3 \leq x \leq 3$, and zero elsewhere. Find the p.d.f if $Y = 2X^2 - 3$. [**Ans.** $F_Y(y) = \{\sqrt{2}\,(3 + y)^{-1/2}/12\}, (-3 < y < 15)$]

5. Suppose that X has the p.d.f.

   (a) $f(x) = \begin{cases} 2x, & 0 < x < 1 \\ 0, & \text{elsewhere} \end{cases}$ (b) $f(x) = \begin{cases} (2/9)\,(x + 1), & -1 < x < 2 \\ 0, & \text{elsewhere} \end{cases}$

   Find the p.d.f. of $Y = 2X + 1$. Find the p.d.f of $Y = X^2$.

6. Let $f_X(x) = \frac{1}{2} e^{-|x|}$, $-\infty < x < \infty$. Find the p.d.f. of (i) $Y = (X)^{1/3}$, (ii) $Z = |X|$.

7. The p.d.f. of a variate X is $f(x) = 1/x^2$, $x \geq 1$; $f(x) = 0$, $x < 1$. Find the distribution of Y defined by $Y = 2X$, if $X \leq 2$; $Y = X^2$, if $X > 2$.

8. Let $f(x) = 2(1 - x)$, $0 < x < 1$; $f(x) = 0$, elsewhere. Find the distribution of Y defined by $Y = 2X$, if $0 < X < \frac{1}{2}$; $Y = 5$, if $X \geq \frac{1}{2}$.

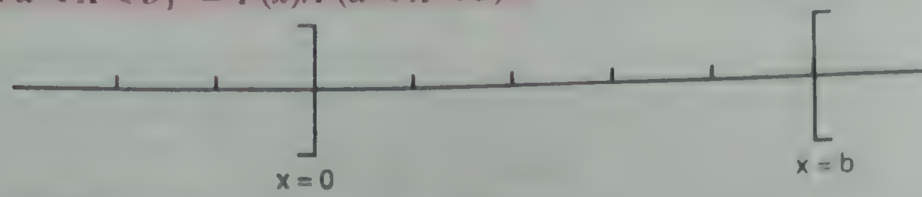
3-74. *Truncated Distributions*

Let X be an integer-valued variate, say $x = 0, 1, 2, 3, \ldots,$ If we *remove* (i.e. truncate) the values of X to the right of $X = b \leq n$, and convert the remaining to a probability function, we get the characterization :

$$P\{X \mid X < b\} = \frac{P(x)}{P(X < b)} = \frac{P(x)}{F(b-1)}, \quad x = 0, 1, 2, \ldots, b - 1.$$

Truncation to the left and to the right of x provides the characterization :

$$P\{X \mid a < X < b\} = P(x)/P(a < X < b).$$



If f is the integrating density of a variate X, then

(i) The p.d.f. of X Truncated to the left of $X = a$ is

$$g(x) = 0, \quad x \le a\,; \quad g(x) = \frac{f(x)}{1-F(a)}, \quad x > a.$$

(ii) The p.d.f. of X, Truncated to the left at $x = a$ and Truncated to the right at $x = b$

$$g(x) = 0, \quad x \le a, \quad x \ge b\,; \quad g(x) = \frac{f(x)}{F(b)-F(a)}, \quad a < x < b.$$

(iii) Some frequently used truncated distributions are bin (n, p), pois(λ), gem(p), $N(\mu, \sigma^2)$, expo(λ), etc.

Example: If, $f(x) = \lambda a^x$, $x = 0, 1, 2, \ldots$, and if X is truncated to the right of $X = k$, such that $P\{0 \mid X \le k-1\}$ is approximately $2P(0)$, find a suitable value of k.

Solution. By Normalization : $1 = \Sigma f(x) = \lambda(1 + a + a^2 + \ldots) = \lambda/(1-a)$, hence $\lambda = (1-a)$. Now

$$P(X \le k-1) = (1-a)\sum_{x=0}^{k-1} a^x = (1-a^k).$$

$$\therefore \qquad P\{X \mid X \le k-1\} = \frac{f(x)}{P(X \le k-1)} = \frac{(1-a)\,a^x}{(1-a^k)}.$$

So $\qquad P(0 \mid X \le k-1) = (1-a)/(1-a^k).$

$\therefore \quad P\{0 \mid X \le k-1\} = 2P(0) \quad \Rightarrow (1-a)/(1-a^k) = 2(1-a) \Rightarrow a^k = \frac{1}{2} \Rightarrow k = \left(\ln \frac{1}{2}\right)/\ln a$

3-80. *Miscellaneous Worked-out Problems*

Example 1. Let $f(x) = x^{a-1}/\{B(a, b)(1+x)^{a+b}\}$; $a > 0$, $b > 0$ $0 > x < \infty$.
Find the distribution of $Y = 1/(1+X)$.

Solution. Here $y = 1/(1+x)$; and obviously, y is decreasing function (since x increases from 0 to ∞) and it decreases from 1 to 0. Now,

$(dy/dx) = -(1+x)^{-2} \quad \Rightarrow \quad dx = -y^{-2}\,dy.$ Also $y = (1+x)^{-1} \quad \Rightarrow \quad x + (1-y)/y.$

$$f(x)\,dx = -\frac{1}{B(a,b)}\cdot\left(\frac{1-y}{y}\right)^{a-1}\cdot y^{a+b}\left(\frac{dy}{y^2}\right) = \frac{y^{b-1}(1-y)^{a-1}}{B(a,b)}\,dy.$$

Reverting the support, hence absorbing negative sign, we recover,

$$g(y) = y^{b-1}(1-y)^{a-1}/B(a, b)\ 0 < y < 1.$$

Remark. The p.d.fs. of X and Y are known as beta-distributions of the second kind and first kind respectively.

***Example* 2.** If $f(x) = e^{-\lambda x}$, $x \geq 0$, $\lambda > 0$, find the distribution of $Y = X/(1 + X)$.

Solution. Here $y = \dfrac{x}{1+x} \Rightarrow x = \dfrac{y}{1-y}; \dfrac{dx}{dy} = \dfrac{1}{(1-y)^2} > 0.$ ...(1)

Since $y' > 0$, it follows that y is an increasing function. Now, from (1)

$x = 0 \Rightarrow y = 0, x \to \infty \Rightarrow y \to 1.$

$$f_Y(y) = \{f_X(x) \mid dx/dy \mid\}_y = \lambda e^{-\lambda y(1-y)}/(1+y)^2, 0 < y < 1.$$

Note. Distribution Function Method is also convenient.

Problems with Solutions Provided at the End of the Text

1*. If the graph of a p.d.f. of a variate X is an equilateral triangle of side-length a, find the value of a.

2*. The p.d.f. of a variate X is $P(X = k) = (1/2)^k$; $k = 1, 2, \ldots$ Find the distribution of $Y = \sin(\pi X/2)$.

3*. Variate X has the p.d.f. $f(x) = 1/\pi$, $-\frac{1}{2}\pi < x \leq \frac{1}{2}\pi$. Find the p.d.f of $Y = \sin X$.

4*. Variate X has the p.d.f. $f(x) = 1$, $0 \leq x \leq 1$. Find the p.d.f. of $Y = e^X$.

Miscellaneous Exercises

1. Suppose X is a variate whose p.d.f. is given by

$p_0 = (x+2)/3, p_3 = (2x+1)/4, p_5 = (1-10x)/12,$ $[p_X = P(X = x)]$.

Determine for what value of x the above is a probability function. **[Ans.** $-\frac{1}{2} \leq x \leq \frac{1}{10}$]

2. Let X be a continuous variate with p.d.f.

$f(x) = ax, 0 \leq x \leq 1, f(x) = 1, 1 \leq x \leq 2, f(x) = -ax + 3a, 2 \leq x \leq 3, f(x) = 0$, elsewhere.

Determine the constant a and show that $P(X \leq 1.5) = \frac{1}{2}$.

3. A discrete variate X has the p.m.f.

$P(X = x) = K/(x+1)(x+2), x = 1, 2, 3, \ldots$ (K : constant)

Determine K, find c.d.f. of X and show that $P(r \leq X \leq s) = 2(s - r + 1)/(r+1)(s+2)$, where r, s are positive integers.

4. Let : $F(x) = 0\, I(x < 0) + x^3\, I(0 \leq x < 1) + 1\, I(x \geq 1)$. Show that $1 - [1 - F(x)]^2$ is a c.p.f. on R.

5. Show that the following functions do not represent c.d.f's. :

(a) $F(x) = \begin{cases} 0, & x < -1 \\ x, & -1 \leq x < 0 \\ 1, & x \geq 0 \end{cases}$; (b) $G(x) = \begin{cases} 0, & x < -1 \\ 1 - x^2, & -1 \leq x < \frac{1}{2} \\ (\frac{1}{2}) + x^2, & \frac{1}{2} \leq x < +1 \\ 1, & x \geq 1 \end{cases}$;

(c) $H(x) = \begin{cases} 0, & x < 0 \\ (\frac{1}{2}) + x/4, & 0 < x < 0. \\ 1, & x > 1 \end{cases}$

6. If the graphs of p.d.f's of variates X and Y are square of side length a and semi-circle of radius b respectively, show that $a = 1$ and $b = (2/\pi)^{1/2}$.
7. Four distinct integers are chosen at random and without replacement from the first 10 positive integers. If the variate X is next to the smallest of these four numbers, show that $f(x) = (x-1)(10-x)(9-x)/420$, $x = 2, 3, \ldots, 8$.
8. Let X have a strictly monotonic, continuous c.d.f. $F(x)$ and let $Y = F(X)$. Show that the distribution of Y is $U(0, 1)$.
9. Let F be a c.d.f. of X. (i) Show that $G(x) = 1 - F(-x^{-}) \; \forall \; x \in R$ is also a c.d.f., (ii) g is the c.d.f. of $-X$.
10. Let X have a strictly monotonic continuous c.d.f. $F(x)$, and let $G(x)$ be strictly monotonic, continuous c.d.f. Show that $G^{-1}(F(x))$ has $G(X)$ as its c.d.f.
11. Variate X has c.d.f. $F(x) = \exp[-e^{-(x-\alpha)/\beta}]$. Show that the distribution of
$$Y = \exp[-(X-\alpha)/\beta] \text{ is } g(y) = e^{-y}, y \ge 0.$$
12. If $f(x) = 2\lambda x e^{-\lambda x^2}$, $x \ge 0$, show that the density of $Y = X^2$ is $g(y) = \lambda e^{-\lambda y}$, $y > 0$.
13. If $f(x) = (1+x)/2$, $-1 \le x \le 1$, show that the density of $Y = X^2$ is $g(y) = \frac{1}{2}\sqrt{y}$, $0 < y < 1$.
14. Variate X has the exponential distribution : $F_X(x) = 1 - e^{-x}, x \ge 0$; $F_X(x) = 0, x < 0$. Let $Y = X$ if $X \le 2$; $Y = 1/X$, if $X > 2$. Find the distribution of Y.

    [**Ans.** $F_Y(t) = 0\, I(t \le 0 + [1 - e^{-t} + e^{-1/t}]\, I(0 < t < \frac{1}{2}) + [1 - e^{-t} + e^{-2}]\, I(\frac{1}{2} \le t < 2) + 1\, I(t > 2)$]
15. The length of time X (measured in days) that a gadget lasts obeys the p.d.f :
$$f(x) = 2e^{-2x}, x > 0 \,;\, f(x) = 0, \text{ elsewhere.}$$
    If the gadget lasts more than 3 days the profit is Rs. 10 ; otherwise it incurs a loss of Rs. 20. Find the distribution of the profit Y. [**Ans.** $P(Y = 10) = e^{-6}, P(Y = -20) = 1 - e^{-6}$]
16. A coin is thrown. If is lands head, a second coin with faces marked 1 and 2 is thrown. If the first coin shows tails then a point is selected at random from the unit interval [0, 1]. If X denotes the number that appears as the final result of the experiment, find c.d.f. of X.
17. A variate X has the c.d.f. 'F' : $F(x) = pH(x) + (1-p).G(x)$, $(0 < p < 1, p \in R)$

    where $H(x) = x, 0 < x < 1$; $H(x) = 1, x \ge 1$, $G(x) = \frac{1}{2}x, 0 < x < 2$; $G(x) = 1, x > 2$.

    Sketch the graph of $F_X(x)$ for $p = \frac{1}{2}$. Give a formula for the p.d.f. of X or the discrete density of X. Also evaluate $P\{X \le \frac{1}{2} \mid X \le 1\}$.
18. A variate X has the c.d.f. :
$$F(x) = \begin{cases} 0, & x < 0 \\ x^2 + 0.2, & 0 \le x < 0.5 \\ x, & 0.5 \le x < 1 \\ 1, & x \ge 1 \end{cases}.$$
    (a) Show that F is partly continuous and partly discrete and sketch the graph of $F(x)$.
    (b) Express $F_X(x)$ in terms of indicator functions.
    (c) Decompose $F_X(x)$ in the form : $pF_c(x) + (1-p)\, F_d(x)$.
    (d) Evaluate $P(0.25 < X < 0.75)$ and $P(0.25 < X < 0.50)$.
19. Show that if $F(t)$ is a c.d.f. and if $G(t) = 1 - F(t)$ satisfies the Cauchy equation $G(x+y) = G(x)\,G(y)$, $\forall\; x, y \ge 0$, then either $F(t)$ is degenerate at zero or $F(t) = 1 - e^{-et}$, $a > 0, t \ge 0$.
    Let X be non-negative r.v. with $P(X \ge 0) = 1$. Deduce that if
    $P(X > x + t\,(X > t) = P(X > x)$, $\forall\; x \ge 0, t \ge 0$, then $F_X(t) = 1 - e^{-at}$, $a > 0, t \ge 0$.

Conduct is wise or foolish only in reference to its results.

Each venture is a new beginning to explore something hidden.

Appendix : Mixed Distributions

C

Definition. Let X be a r.v. with distribution function F. If $F(x)$ can be decomposed as

$$F(x) = pF_d(x) + (1-p)\,F_{ac}(x), \quad 0 < p < 1 \qquad \ldots(1)$$

where F_d is the c.d.f. of some discrete variate and F_{ac} is the c.d.f. of some continuous variate (F_{ac} is absolutely continuous), then F is called *mixed distribution function.* The definition implies that F is partly continuous and partly discrete, i.e. F is neither continuous nor discrete.

Comments. The density function $f(x)$ corresponding to $F(x)$ can be defined by

$$f(x) = p\,f_d(x) + (1-p)\,f_{ac}(x) \quad [f_d : \text{discrete density}, f_{ac} : \text{abs. continuous density}] \ldots(2)$$

The definition (2) is more involved; hence we would prefer using (1).

Remarks. In general, it can be proved that any distribution function F can be decomposed in the form

$$F(x) = p_1 F_d(X) + p_2 F_{ac}(x) + p_3 F_s(x)$$

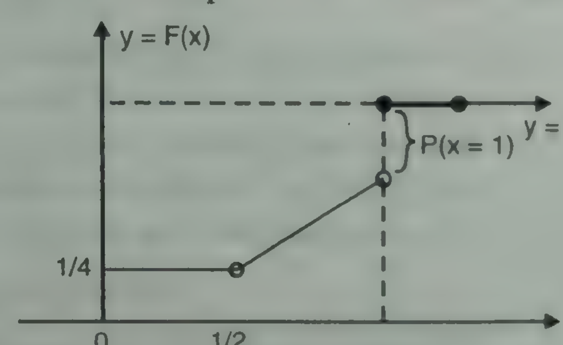
where $0 \le p_1, p_2, p_3 \le 1$; $p_1 + p_2 + p_3 = 1$.

F_d is the c.d.f. of some discrete variate, F_{ac} is the c.d.f. of some absolutely continuous variate and F_s is the singular distribution function.

The proof of this result is beyond the scope of this book. We now present two illustrations for mixed distributions.

***Example* C-1.** F is a c.d.f. on R defined as

$$F(x) = \begin{cases} 0, & x < 0 \\ \frac{1}{4}, & 0 < x < \frac{1}{2} \\ x - (\frac{1}{4}), & \frac{1}{2} \le x < 1 \\ 1, & x > 1 \end{cases}.$$



Show that it is neither discrete nor continuous.

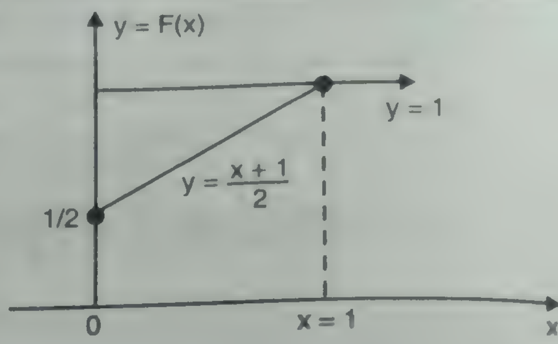
Solution. The graph of F is shown in the figure above ; since it is not a staircase (step function), it is not discrete. It is neither continuous because it has jumps (discontinuties) at $x = 0$ and at $x = 1$. Obviously.

$P(X = 1) = F(1) - F(1-0) = 1 - (1 - \frac{1}{4}) = \frac{1}{4}$. $P(X = 0) = F(0) - F(0-) = \frac{1}{4} - 0 = \frac{1}{4}$.

In fact $F(x) = \frac{1}{2} F_d(x) + \frac{1}{2} F_c(x)$, where

$$F_d(x) = \begin{cases} 0, & x < 0 \\ \frac{1}{2}, & 0 \le x < 1 \\ 1, & x \ge 1 \end{cases}; \qquad F_c(x) = \begin{cases} 0, & x \le \frac{1}{2} \\ 2x - 1, & \frac{1}{2} \le x \le 1 \\ 1, & x \ge 1 \end{cases}.$$

***Example* C-2.** A random variable X has the c.d.f. 'F'



$$F(x) = \begin{cases} 0, & x<0 \\ \frac{1}{2}, & x=0 \\ (\frac{1}{2})+(\frac{x}{2}), & 0<x<1 \\ 1 & x \ge 1 \end{cases}$$

Show that f is neither continuous nor discontinuous.

Also evaluate $P(0 \le X \le \frac{1}{2})$.

Solution. The graph of F is shown in the figure above. Since it is not a stair case (step function), it is not discrete. It is neither continuous because it has jump (discontinuity) at $x = 0$. Now

$$P(X=0) = F(0) - F(0-0) = \tfrac{1}{2} - 0 = \tfrac{1}{2}.$$

We can also decompose : $F(x) = \frac{1}{2} F_d(x) + \frac{1}{2} F_c(x)$,

where $$F_d(x) = \begin{cases} 0, & x \le 0 \\ x, & 0<x<1; \\ 1, & x \ge 1 \end{cases} \qquad F_c(x) = \begin{cases} 0, & x<0 \\ 1, & x \ge 1 \end{cases};$$

$$P(0 \le X \le \tfrac{1}{2}) = F\left(\tfrac{1}{2}\right) - F(0) + P(X=0) = \left(\tfrac{1}{2}+\tfrac{1}{4}\right) - \tfrac{1}{2} + \tfrac{1}{2} = \tfrac{3}{4}.$$

Note. $F_d(x)$ is the d.f. of the variate Degenerate at $x = 0$.

Some Exercises

Note. **Indicator Form :** $F(x) = g(x)\, I(a<x\le b) + h(x)\, I(c<x\le d)$ means :

$$F(x) = g(x), a<x\le b \,;\, F(x) = h(x), c<x\le d, \text{ etc.}$$

1. A random variable X has the c.d.f. F

$F(x) = 1 - pe^{-\lambda x}, x>0, \lambda>0, 0<p<1 \,;\, F(x) = 0, x \le 0.$

Decompose $F(x)$ in the form $pF_c(x) + (1-p)\,F_d(x)$, where F_c is the d.f. of a continuous variate and F_d is the d.f. of a discrete variate. Find $P(X>2)$, $P(X\le 0)$ and $P(X=0)$.

2. A random variable X has the c.d.f. 'F'

$$F(x) = 0\, I(x<0) + \tfrac{1}{2}(x+1)\, I(0 \le x<1) + 1\, I(x \ge 1).$$

Show that F is neither continuous nor discrete. Prove that

$$P(X=0) = \tfrac{1}{2} \text{ and } P(-3<X\le \tfrac{1}{2}) = \tfrac{3}{4}.$$

3. A random variable X has the c.d.f. 'F'

$$F(x) = 0\, I(x<0) + \tfrac{1}{2}\, I(0 \le x<\tfrac{1}{2}) + x\, I(\tfrac{1}{2} \le x \le 1) + 1\, I(x>1).$$

Show that F is neither continuous nor discrete. Decompose $F(x)$ in the form $pF_c(x) + (1-p)\,F_d(x), 0<p<1$. Show further : $P(\frac{1}{2} \le X \le 1) = \frac{1}{2}$.

4. Let the distribution function, 'F' in the indicator form be :

$$F(x) = 0\, I(x<-1) + \tfrac{1}{4}(x+2)\, I(-1 \le x<1) + 1\, I(x \ge 1).$$

Sketch the graph of $F(x)$ and write it as $pF_c(x) + (1-p)\,F_d(x), 0<p<1$.

Evaluate : $P(-\frac{1}{2} < X \le -\frac{1}{3})$, $P(X = 0)$, $P(2 < x \le 3)$.

5. A random variable X has the c.d.f. F

$F(x) = 0,\ x < 0;\ \ F(x) = 1 - (\frac{1}{2})^{x+1} - (\frac{1}{2})^{[x]+1},\ \ x \ge 0.$

Show that F is neither continuous nor discrete. Sketch the graph of $F(x)$ and decompose it as $pF_c(x) + (1-p)\,F_d(x)$. Also show that,

$$P(1 < X \le 3.5) = (7/16) - (1/\sqrt{2})^9;\ \ P(1.5 \le X < 4.5) = (7/32) - (1/\sqrt{2})^{11} + (1/\sqrt{2})^5;$$

$$P(1 < X < 4) = 13/32.$$

6. A random variable X has the mixed c.d.f.

$$F(x) = 0\ I(x < -1) + k(x+2)\,I(-1 \le x < 1) + 1\,I(x \ge 1).$$

Given that $P(X = 1) = \frac{1}{3}$, find k. Sketch the graph of $F(x)$ and decompose it as $pF_c(x) + (1-p)$ $F_d(x)$. Determine : $P(X = -1)$, $P(-1 \le X < 1)$, $P(|X| < 1)$, $P(X \ge \frac{1}{2} \mid X > \frac{1}{3})$.

7. A random variable X, as the c.d.f. F in the indicator form :

$F(x) = 0\,I(x < 0) + \frac{1}{3}x\,I(0 \le x < 1) + \frac{1}{2}x\ I(1 \le x < 2) + 1\,I(x \ge 2).$

Sketch the graph of $F(x)$ and decompose it as $pF_c(x) + (1-p)\,F_d(x)$. Also evaluate

$P(\frac{1}{2} \le X \le \frac{3}{2})$, $P(\frac{1}{2} \le X \le 1)$. $P(\frac{1}{2} \le X < 1)$, $P(1 < X < 2)$.

8. A variate X has the c.d.f. F, expressed in indicator form :

$F(x) = 0\,I\,(x < 0) + \ xI(0 \le x < \frac{1}{2}) + \frac{3}{4}\,I(\frac{1}{2} \le x < 1)\ + [(x/8) + (6/8)]\,I(1 \le x < 2) + 1\,I(x \ge 2).$

Show that F is neither continuous nor discrete. Sketch the graph of $F(x)$ and decompose it as $pF_c(x) + (1-p)\,F_d(x)$. Also show that,

$P(X = \frac{1}{2}) = \frac{1}{4}$; $P(0.5 < X \le 1.5) = \frac{3}{16}$; $P(\frac{1}{4} < X \le \frac{3}{4}) = \frac{1}{2}$; $P(X > \frac{3}{4} \mid X \le 1.5) = \frac{1}{5}$.

And of all axioms this shall win the prize
It's better to be fortunate than wise. (John Webster)

The power of imagination makes us infinite. (John Muir)

Jointly Distributed Random Variables

4

4-00. *Introduction*

When the outcome of a random experiment can be characterized in more than one way, the probability density is a function of more than one variate. For example, when a card is drawn from an ordinary deck, it may be characterized according to its suit and to its denomination. Let X be a variate that assumes the values 1, 2, 3, 4 which correspond to suit in some order (say Clubs, Diamonds, Hearts, Spades) and Y be a variate that assumes the values 1, 2, ..., 13 which correspond to the denominations : Ace, 2, 3, ..., 10, J, Q, K. Then (X, Y) is a 2-dimensional variate. The probability of drawing a particular card will be denoted by $f(x, y)$ and if each card is equi-probable of being drawn, the density of (X, Y) is

$$f(x, y) = 1/52 \; ; \; 1 \le x \le 4, \; 1 \le y \le 13.$$

Trials whose outcomes can be characterized by two (three) variates give rise to bivariate (tri-variate) distributions, etc.

Extensions to n-variate distributions are fairly straight-forward.

4-10. *Two Dimensional Random Vector (2-dim r.v.)*

Definition. Let $(\Omega, \mathcal{F}, P)$ be probability pace and consider two functions $X : \Omega \to R$ and $Y : \Omega \to R$. Then vector (X, Y) is a function :

$$(X, Y) : \Omega \to R \times R \text{ defined by } (X, Y)(\omega) = [X(\omega), Y(\omega)], \; \forall \; \omega \in \Omega.$$

A vector (X, Y) is called a 2-dimensional random vector iff $\{\omega : X(\omega) \le x, Y(\omega) \le y\} \in \mathcal{F}$, for all real numbers x and y.

Note. n-dimensional random vector $(X_1, X_2, ..., X_n)$ is similarly defined :

$(X_1, X_2, ..., X_n) : \Omega \to R^n$ is a r.v. iff $\{X_1 \le x_1, ..., X_n \le x_n\} \in \mathcal{F}, \; \forall \; x_i$

where, as usual, we abbreviate : $\{\omega : X(\omega) \le x, Y(\omega) \le y\} \equiv \{X \le x, Y \le y\}$, etc.

Theorem. The vector (X, Y) is a random vector on a probability space $(\Omega, \mathcal{F}, P)$ iff X and Y are random variables on $(\Omega, \mathcal{F}, P)$.

Proof. Let (X, Y) be a random vector on the probability space $(\Omega, \mathcal{F}, P)$. Then, by definition,

$$\{X \le x, Y \le y\} \in \mathcal{F}, \; \forall \; x, y \in \mathbb{R}.$$

In particular, $\{X \le x, Y < \infty\} \in \mathcal{F}$. But $\{X \le x, Y < \infty\} = \{X \le x\} \cap \{Y < \infty\} = \{X \le x\} \cap \Omega = \{X \le x\}$

$\therefore \quad \{X \le x\} \in \mathcal{F}, \; \forall \; x \in R \;\; \Rightarrow \;\; x$ is a r.v. on $(\Omega, \mathcal{F}, P)$.

Similarly, Y is also a r.v. on $(\Omega, \mathcal{F}, P)$.

Conversely. Let X and Y be variates on $(\Omega, \mathcal{F}, P)$ and x, y be any two real numbers. Then, by definition, $\{X \le x\} \in \mathcal{F}, \forall\, x \in \mathbb{R}$; $\{Y \le y\} \in \mathcal{F}, \forall\, y \in \mathbb{R}$. By the nature of σ-field $\mathcal{F}$:

$$\{X \le x\} \cap \{Y \le y\} \in \mathcal{F}, \text{ i.e. } \{X \le x, Y \le y\} \in \mathcal{F}, \forall\, x, y \in \mathbb{R}.$$

This proves that (X, Y) is a random vector.

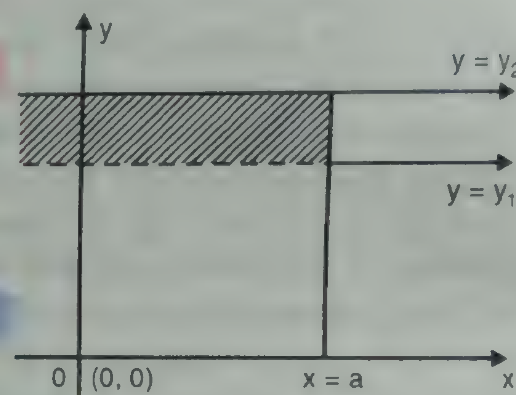
4-11. *Joint Distribution Function and its Properties*

Let (X, Y) be a random vector on the probability space $(\Omega, \mathcal{F}, P)$. The joint c.d.f of X and Y is denoted by $F_{Y,Y}$ and is defined by

$$F_{X,Y}(x, y) = P(X \le x, Y \le y), \quad \forall\, x, y \in \mathbb{R}.$$

Note. The individual c.d.fs. of X and Y shall be denoted by F_X (or F_1), and F_Y (or F_2) and $F_{X,Y}$ shall be simply written as F.

A joint. c.d.f. of two variates has the following properties :



1. ***Non-negativity and Boundedness :*** $0 \le F(x, y) \le 1$, $\forall\, x, y \in \mathbb{R}$.

Proof. $F(x, y) = P(X \le x, Y \le y), \ \forall\, x, y \in \mathbb{R}$.

As probability measure is non-negative and $0 \le P(X \le x, Y \le y) \le 1$, hence (1) follows :

2. ***Monotonicity :*** The c.d.f 'F' is monotonically non-decreasing function in each of the individual variables, i.e.

(i) $F(a, y_2) \le F(a, y_1)$, if $y_2 \ge y_1$,

(ii) $F(x_2, b) \ge F(x_1, b)$, if $x_2 \ge x_1$.

Proof. Observe the equivalent events

$$\{X \le a, Y \le y_2\} = \{X \le a, Y \le y_1\} \uplus \{X \le a, y_1 < Y \le y_2\}.$$

Taking probability of both sides, using finite-additivity :

$$P\{X \le a, Y \le y_2\} = P\{X \le a, Y \le y_1\} + P\{X \le a, y_1 < Y \le y_2\}$$

or $P\{X \le a, y_1 < Y \le y_2\} = F(a, y_2) - F(a, y_1)$. ...(1)

Since 'P' is non-negative measure, this readily gives : $F(a, y_2) \ge F(a, y_1)$.

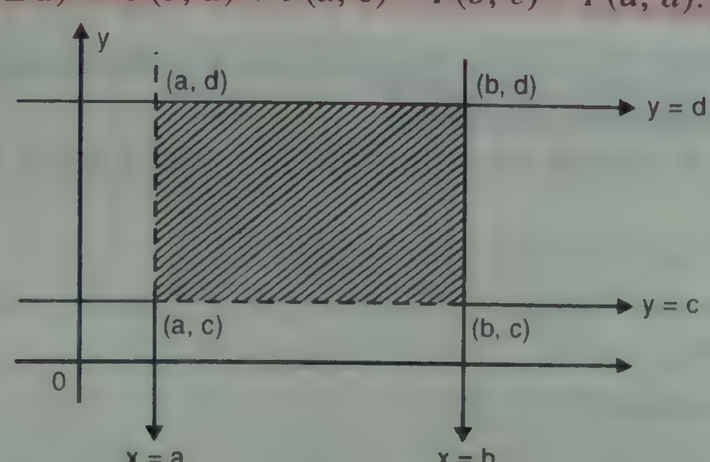
We can similarly establish (ii).

Cor. $P(X \le a, c < Y \le d) = F(a, d) - F(a, c) \ (d > c)$ [by (1)] ...(2)

3. ***Rectangular Rule :*** Let a, b, c, d be any real numbers with $a < b$ and $c < d$. Then,

$$P(a < X \le b, c < Y \le d) = F(b, d) + F(a, c) - F(b, c) - F(a, d).$$

$\left({}^{a}_{c}\times{}^{b}_{d}\right)$



Proof. We define the events : $A = \{X \le a\}$, $B = \{X \le b\}$, $C = \{Y \le c\}$, $D = \{Y \le d\}$.

$\therefore \quad P\{a < X \le b, c < Y \le d\} = P\{(B - A) \cap (D - C)\} = P\{B \cap (D - C) - A \cap (D - C)\}$

$$= P\{B \cap (D - C) - P\{A \cap (D - C)\} \qquad [\because A \subseteq B]$$

$$= P\{B \cap D - B \cap C\} - P\{A \cap D - A \cap C\}$$

$$= [P(B \cap D) - P(B \cap C)] - [P(A \cap D) - P(A \cap D)] \qquad (\because C \subseteq D)$$

$$= P(X \le b, Y \le d) - P(X \le b, Y \le c) - P(X \le a, Y \le d) + P(X \le a, Y \le c)$$

$$= F(b, d) - F(b, c) - F(a, d) + F(a, c).$$

4. ***Individual Limits*** : (i) $\lim\limits_{n \to \infty} F(x, y) = F(-\infty, y) = 0$; (ii) $\lim\limits_{y \to -\infty} F(x, y) = F(x, -\infty) = 0$.

Proof. (i) Let $\langle x_n \rangle$ be any decreasing sequence of real numbers such that $\lim x_n = -\infty$ as $n \to \infty$. Then the sequence of sets $A_n = \{X \le x_n, Y \le y\}$ is a decreasing sequence and hence

$$\lim_{n \to \infty} A_n = \bigcap_{n=1}^{\infty} A_n = \varnothing.$$

$$\therefore \qquad \lim_{n \to \infty} F(x_n, y) = \lim_{n \to \infty} P(X \le x_n, Y \le y) = \lim_{n \to \infty} P(A_n) = P(\lim_{n \to \infty} A_n) = P(\varnothing) = 0$$

as P is a continuous measure. Thus, by the notation set-out $F(-\infty, y) = 0$. Similarly, $F(x, -\infty) = 0$.

5. ***Double Limit*** : $\lim\limits_{x \to \infty, y \to \infty} F(x, y) = 1$, i.e. $F(\infty, \infty) = 1$ (Formally).

Proof. Let $\langle x_n \rangle$ and $\langle y_n \rangle$ be two increasing sequences of real numbers such that

$$\lim_{n \to \infty} x_n = +\infty, \lim_{n \to \infty} y = +\infty$$

Then, the sequence of set $A_n = \{X \le x_n, Y \le y_n\}$ is an increasing sequence and so

$$\lim_{n \to \infty} A_n = \bigcup_{n=1}^{\infty} A_n = \Omega.$$

$$\therefore \qquad \lim_{n \to \infty} F(x_n, y_n) = \lim_{n \to \infty} P(X \le x_n, Y \le y_n) = \lim_{n \to \infty} P(A_n) = P(\lim_{n \to \infty} A_n) = P(\Omega) = 1.$$

[P is a continuous function]

Thus, the stated double limit (5) holds.

Remark. We don't claim $F(\infty, y) = 1$ or $F(x, \infty) = 1$.

6. ***Individual Continuity*** : F is continuous from the right in each of its individual variables, i.e.

(i) $\lim\limits_{x \to a+} F(x, y) = F(a, y)$ (ii) $\lim\limits_{y \to b+} F(x, y) = F(x, b)$.

Proof. (i) Let $\langle x_n \rangle$ be an increasing sequence of real numbers such that $x_n > a$ and $\lim x_n = a$. Then, the sequence of sets $A_n = \{X \le x_n, Y \le y\}$ is a decreasing sequence and so

$$\lim_{n \to \infty} A_n = \bigcap_{n=1}^{\infty} A_n = \{X \le a, Y \le y\}.$$

Using continuity of measure P we have

$$\lim_{n \to \infty} F(x_n, y) = \lim_{n \to \infty} P(X \le x_n, Y \le y) = \lim_{n \to \infty} P(A_n) = P(\lim_{n \to \infty} A_n) = P(X \le a, Y \le y) = F(a, y).$$

Thus $\lim_{x \to a^+} F(x, y) = F(a, y)$. Proof of (ii) is exactly similar.

Consistency Relationship :

$P\{X \le x$, no condition on $Y\} = F_X(x)$, P {No condition on X, $Y \le y\} = F_Y(y)$.

In particular, $F_{X,Y}(x, \infty) = F_X(x)$, $\quad F_{X,Y}(\infty, y) = F_Y(y)$.

7. $P(X = a, Y = b) = F(a, b) + F(a^-, b^-) - F(a, b^-) - F(a^-, b)$ where $F(a^+, b) = \lim_{x \to a^+} F(x, b)$, etc.

Proof. Let us consider the sequence of sets :

$$A_n = (a - n)^{-1} < X \le a, b - n^{-1} < Y \le b$$

Then, this sequence $\langle A_n \rangle$ of sets is a decreasing sequence and so

$$\lim_{n \to \infty} A_n = \bigcap_{n=1}^{\infty} A_n = \{X = a, Y = b\}.$$

Using continuity of measure P and the Rectangle Rule, we have

$$\begin{aligned} P(X = a, Y = b) &= P\left(\lim_{n \to \infty} A_n\right) = \lim_{n \to \infty} P(A_n) = \lim_{n \to \infty} P\left(a - \frac{1}{n} < X \le a, b - \frac{1}{n} < Y \le b\right) \\ &= \lim_{n \to \infty} \left[F\left(a - \frac{1}{n}, b - \frac{1}{n}\right) + F(a, b) - F\left(a - \frac{1}{n}, b\right) - F\left(a, b - \frac{1}{n}\right)\right] \\ &= F(a^-, b^-) + F(a, b) - F(a^-, b) - F(a, b^-) \end{aligned}$$

as the sequences $\langle a - n^{-1} \rangle$ and $\langle b - n^{-1} \rangle$ are increasing and satisfy $a - n^{-1} < a$, $b - n^{-1} < b$;

$$\lim (a - n^{-1}) = a, \; \lim (b - n^{-1}) = b.$$

8. ***Survival Relation*** : $F_{X,Y}(a, b) - F_X(a) - F_Y(b) = P(X > a, Y > b) - 1$.

Proof. Define the *survival* functions $A = \{X > a\}$, $B = \{Y > b\}$ so that $\overline{A} = \{X \le a\}$, $\overline{B} = \{Y \le b\}$.

$$\begin{aligned} F(a, b) = P(X \le a, Y \le b) &= P(\overline{A}\,\overline{B}) = 1 - P(A \cup B) = 1 - P(A) - P(B) + P(AB) \\ &= P(\overline{A}) + P(\overline{B}) + P(AB) - 1 = F_X(a) + F_Y(b) + P(X > a, Y > b) - 1. \end{aligned}$$

Remarks. If F is x-continuous, then $F(a, b) = F(a^-, b)$ as also, $F(a, b^-) = F(a^-, b^-)$ and thus $P(X = a, Y = b) = 0$. Similarly if F is y-continuous, $P(X = a, Y = b) = 0$. Thus, to locate the points (x, y) where $P(X = x, Y = y) > 0$, we need consider only those points where F is discontinuous in *both* variables.

Summary : Some Properties of Bivariate c.d.f. $\quad [F_{XY}(x, y) \equiv F(x, y)]$

1. $0 \le F(x, y) \le 1$
2. If $x_1 \le x_2$ and $y_1 \le y_2$, then
   $F(x_1, y_1) \le F(x_2, y_1) \le F(x_2, y_2)$; $\quad F(x_1, y_1) \le F(x_1, y_2) \le F(x_2, y_2)$
3. $\lim_{\substack{x \to \infty \\ y \to \infty}} F(x, y) = F(\infty, \infty) = 1$.
4. $\lim_{x \to -\infty} F(x, y) = F(-\infty, y) = 0$; $\quad \lim_{y \to -\infty} F(x, y) = F(x, -\infty) = 0$
5. $\lim_{x \to a^+} F(x, y) = F(a^+, y) = F(a, y)$; $\quad \lim_{y \to b^+} F(x, y) = F(x, b^+) = F(x, b)$
6. $F\{a < X \le b, Y \le y\} = F(b, y) - F(a, y)$; $\quad P\{X \le x, c < y \le d\} = F(x, d) - F(x, c)$
7. If $a \le b, c \le d$, then
   $P\{a < X \le b, c < Y \le d\} = F(b, d) + F(a, c) - F(a, d) - F(b, c)$.

4-12. Worked-out Problems

***Example* 1.** Functions $G(x, y)$ and $H(x, y)$ of two variates X, Y are defined by (i) $G(x, y) = 1$, if $x + y \geq 0$; $G(x, y) = 0$ if $x + y < 0$. (ii) $H(x, y) = 1$, if $x \geq y$, $H(x, y) = 0$, if $x < y$. Examine whether G and H can be the distribution functions of some 2-dimensional distributions.

Solution. (i) Function $G(x, y)$ is a c.d.f. if Rectangle Rule holds. Now

$$P(a < X \leq b, c < Y \leq d) = G(b, d) + G(a, c) - G(a, d) - G(b, c).$$

$$P(-1 < X \leq 0, 0 < Y \leq 1) = G(0, 1) + G(-1, 0) - G(-1, 1) - G(0, 0) = 1 + 0 - 1 - 1 = -1.$$

Since the probability measure turns out to be negative, the function G cannot be a c.d.f.

(ii) We choose $a < c < b < d$ and use Rectangle Rule :

$$P(a < X \leq b, c < Y \leq d) = H(a, c) + H(b, d) - H(b, c) - H(a, d) = 0 + 0 - 1 - 0 = -1.$$

Since the probability measure is to be non-negative, $H(x, y)$ in not a c.d.f.

***Example* 2.** If X and Y have a joint c.d.f. F, show that

$$F_1(x) + F_2(y) - 1 \leq F(x, y) \leq [F_1(x)\, F_2(y)]^{1/2}, \forall\, x, y.$$

Solution. Let $A = \{X \leq x\}$, $B = \{Y \leq y\}$; then $A \cap B = \{X \leq x, Y \leq y\}$.

$$A \cap B \subseteq A \Rightarrow P(AB) \leq P(A);\text{ like wise } P(AB) \leq P(B)\text{, hence}$$

$$[P(AB)]^2 \leq P(A)\, P(B) \Rightarrow [F(x, y)]^2 \leq F_1(x) \,.\, F_2(y) \Rightarrow F(x, y) \leq [F_1(x) \,.\, F_2(y)]^{1/2} \quad \ldots(1)$$

$$P(A \cup B) = P(A)\, P(B) - P(AB) \Rightarrow P(AB) \geq P(A) + P(B) - 1 \quad [\because P(A \cup B) \leq 1] \ldots(2)$$

$$\therefore \quad F(x, y) \geq F_1(x) + F_2(y) - 1.$$

Combining (1) and (2) we get $F_1(x) + F_2(y) - 1 \leq F(x, y) \leq [F_1(x)\, F_2(y)]^{1/2}$.

Problems with Solutions Provided at the End of the Text

1*. The joint c.d.f. of X and Y is given by

$$F(x, y) = \begin{cases} 0, & x < -2 \quad \text{or} \quad y < -5 \\ \frac{1}{4}, & -2 \leq x < 2, \text{ and } -5 \leq y < 3 \\ \cancel{\frac{1}{2}}, & \cancel{-2 \leq x < 2, \text{ and } -5 \leq y < 3} \\ \frac{1}{2}, & -2 \leq x < 2, \text{ and } \quad y \geq 3 \\ 1 & \quad x \geq 2, \text{ and } \quad y \geq 3 \end{cases}$$

Find $P(X = -2, Y = -5)$, $P(X = -2, Y = 3)$, $P(X = 2, Y = -5)$, $P(X = 2, Y = 3)$.

2*. Show that the bivariate function

$G(x, y) = 1 - e^{-2x}$, $x > 0$, $y \geq 1$; $G(x, y) = \frac{1}{2}(1 - e^{-x})$, $x > 0$, $0 \leq y < 1$ is not a distribution function although its marginals are distribution functions.

3*. Find the chance that a point (X, Y) hits the region : $1 \leq X \leq 2$, $1 \leq y \leq 2$ if

$F(x, y) = 1 - a^{-x^2} - a^{-2y^2} + a^{-x^2-2y^2}, x \geq 0, y \geq 0, (a > 0)$; $F(x, y) = 0$, $x < 0$, $y < 0$.

4*. Give examples of *different* bivariate distribution functions which posses the *same* marginal distribution functions.

4-20. Joint Discrete Distribution Function

Let (X, Y) be a random vector with joint c.d.f. 'F'. Recall

$$P(X = x, Y = y) = F(x, y) + F(x^-, y^-) - F(x, y^-) - F(x^-, y^-)$$

Since F is a real-valued bounded function, and is monotonic in each variable, so F has finite or countably infinite number of points for which $P(X = x, Y = y) > 0$.

Definition. The joint c.d.f. of X and Y is said to be *discrete* if there exists a non-negative function p such that p vanishes everywhere except at a finite or countably infinite number of points in the plane and at such points (x, y), $p(x, y) = P(X = x, Y = y)$, $x, y \in R$.

Theorem. If the set of points where p does not vanish is denoted by $\{(x_i, y_j) : i, j = 1, 2, 3, \ldots\}$, then the following result holds :

(i) $p(x_i, y_j) \geq 0 \quad \forall\, i, j$ (ii) $F(x, y) = \sum\limits_{x_i \leq x} \sum\limits_{y_j \leq y} p(x_i, y_j)$ (iii) $\sum\limits_{i=1}^{\infty} \sum\limits_{j=1}^{\infty} p(x_i, y_j) = 1.$

Proof. (i) Since $p(x_i, y_j) = P(X = x_i, Y = y_j)$, so $p(x_i, y_j) \geq 0$, since probability measure is non-negative

(ii) $F(x, y) = P(X \leq x, Y \leq y) = \sum\limits_{x_i \leq x} P(X = x_i, Y \leq y)$ [P is countable additive]

$$= \sum_{x_i \leq x} \left[\sum_{y_j \leq y} P(X = x_i, Y = y_j) \right] = \sum_{x_i \leq x} \sum_{y_j \leq x} p(x_i, y_j) \text{ [countable additivity of } P]$$

(iii) $1 = \lim\limits_{x \to \infty, y \to \infty} F(x, y) = \lim\limits_{x \to \infty, y \to \infty} \sum\limits_{x_i \leq x} \sum\limits_{y_j \leq y} p(x_i, y_j) = \sum\limits_{x_i < \infty} \sum\limits_{y_j < \infty} p(x_i, y_j) = \sum\limits_{i=1}^{\infty} \sum\limits_{j=1}^{\infty} p(x_i, y_j).$

Definition. Let X and Y have a joint *discrete* distribution. A function p which does not vanish on the set $\{(x_i, y_j) ; i, j = 1, 2, \ldots\}$ and satisfies the properties :

(i) $p(x_i, y_j) \geq 0$, $i, j = 1, 2, 3, \ldots$ (ii) $\sum\limits_{i=1}^{\infty} \sum\limits_{j=1}^{\infty} p(x_i, y_j) = 1$

is called **joint probability (mass) function** of X and Y or simply the **joint probability function.** Often we use the term, *discrete density* for brevity.

4-21. Individual or Marginal Probability Functions

Let X and Y be two jointly distributed variates with joint discrete density $p(x, y)$. By theorem (§4-10), the individual variates X and Y themselves are random variables. The individual distribution of X and Y are called *marginal* distribution of X and Y.

(1) The marginal probability function for X is denoted by $p_X(x)$ or $p(x)$ and is given by

$$p(x) = P(X = x) = \textstyle\sum_y P(X = x, Y = y) = \sum_y p(x, y).$$

(2) The marginal probability function for Y is denoted by $p_Y(y)$ or $p(y)$ and is given by

$$p(y) = P(Y = y) = \textstyle\sum_x P(X = x, Y = y) = \sum_x p(x, y).$$

Note. It is often convenient to display the probability function of a bivariate distribution in a rectangular array, in which the row totals and column totals provide the marginal probability functions of X and Y respectively. Thus,

Bivariate Probability Distribution Table

$x \downarrow$ \ $y \rightarrow$	y_1	y_2	$\cdots$	y_j	$\cdots$	y_m	$\cdots$	$p(X = x_i)$
x_1	p_{11}	p_{12}	$\cdots$	p_{1j}	$\cdots$	p_{1m}	$\cdots$	$p(x_1)$
x_2	p_{21}	p_{22}	$\cdots$	p_{2j}	$\cdots$	p_{2m}	$\cdots$	$p(x_2)$
$\vdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
x_i	p_{i1}	p_{i2}	$\cdots$	p_{ij}	$\cdots$	p_{im}	$\cdots$	$p(x_i)$
$\vdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
x_n	p_{ni}	p_{n2}	$\cdots$	p_{nj}	$\cdots$	p_{nm}	$\cdots$	$p(x_n)$
$\vdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
$P(y = y_j)$	$p(y_1)$	$p(y_2)$	$\cdots$	$p(y_j)$	$\cdots$	$p(y_m)$	$\cdots$	1

Here, $p_{ij} = P(X = x_i, Y = y_j)$, $\quad p(x_i) = \Sigma_j\, p_{ij}$, $\quad p(y_j) = \Sigma_i\, p_{ij}$.

4-22. Conditional Distributions

Recall that if A and B are two events with $P(B) > 0$, then $P(A \mid B) = P(AB)/P(B)$. Letting $A = \{X \le x\}$, we get

$$P\{X \le x \mid B\} = P\{(X \le x) \cap B\}/P(B).$$

This is called the *conditional distribution function* of variate X given that the event B occurs.

Conditional Probability Functions (Cond. p.f.)

Let X and Y have a joint discrete distribution with associated probability function p. Let the possible values of X be $\{x_1, x_2, \ldots, x_i \ldots\}$ and those of Y be $\{y_1, y_2, \ldots, y_j, \ldots\}$ respectively.

The *conditional probability function of* X, given $Y = y_j$ notated by $P_{X \mid y_j}(x_i \mid y_j)$, is defined by

$$\begin{aligned} P_{X \mid y_j}(x_i \mid y_j) &= P(x_i, y_j)/P_y(y_j),\ i = 1, 2, 3, \ldots \\ &= 0, \text{ if } P_Y(y_j) = 0. \end{aligned}$$

The *conditional probability function of* Y, given $X = x_i$, notated $P_{X \mid x_i}(y_j \mid x_i)$ is defined by

$$\begin{aligned} P_{Y \mid x_i}(y_j \mid x_i) &= P(x_i\,, y_j)/P_X(x_j),\ j = 1, 2, 3, \ldots \\ &= 0, \text{ if } P_X(x_j) = 0. \end{aligned}$$

Recall : $\quad P(x_i, y_j) \equiv P(X = x_i, Y = y_j)$, $P(Y = y_j) = P_Y(y_j)$, and $P(X = x_i) = P_x(x_i)$.

Theorem. Conditional probability function of X, given $Y = y_j$ is a bona-fide discrete p.f.

Proof. (i) $P\{X = x_i \mid Y = y_1\} = P(x_i, y_j)/P(y_j) \ge 0$, since probability is non-negative measure.

(ii) $$\sum_{i=1}^{\infty} P(X = x_i \mid Y = y_j) = \sum_{i=1}^{\infty} \frac{P(x_i, y_j)}{P_Y(y_j)} = \frac{P_Y(y_j)}{P_Y(y_j)} = 1$$

where we have used marginal density of $Y = y_j$ [Multi-stage p-Rule]

Since $P_{X \mid y_j}(x_i, y_j) \ge 0$ and $\Sigma_i\, P_{X \mid y_i}(x_i, y_j) = 1$, the theorem is proved.

4-30. Independent and Dependent Random Variables

***Definition* 1.** Two random variables X and Y are called **independent** if for *every* pair of real numbers x and y, the two events $\{X \le x\}$ and $\{Y \le y\}$ are independent ; i.e.

$$P\{X \le x, Y \le y\} = P\{X \le x\}\, P\{Y \le y\} \qquad \text{...(1)}$$

Condition (1) in terms of distribution functions is : $F(x, y) = F_X(x)\, F_Y(y)$ ...(2)

Also : $f(x, y) = f_X(x) f_Y(y)$, [if densities exists] ...(3)

Conversely, if (2) or (3) is true, then (1) follows.

***Definition* 2.** Two variates X and Y are called **independent**, if for every choice of real numbers $a, b, c, d, (a < b, c < d)$

$$P(a < X \le b, c < Y \le d) = P(a < X \le b)\, P(c < Y \le d).$$

Definitions 1 and 2 are really equivalent.

Dependent variates. Variates which are not independent are called *dependent* random variables.

Theorem. Let X and Y be independent variates and suppose that $G(X) = g$, $H(Y) = h$. Then G, H are also independent.

We assume that G and H are defined on the respective ranges of X and Y. Then

$$P\{G(X) = g, H(Y) = h\} = \Sigma_A P\{X = x, Y = y\}, \quad A = \{x \in G^{-1}(g), y \in H^{-1}(h)\}$$
$$= \Sigma_A P(X = x)\,.\,P(Y = y)$$
$$= P(G(X) = g\}\,.\,P\{H(Y) = h\}.$$

Comments. The converse result of this theorem is not true, i.e. if $g(X)$ and $h(X)$ are independent, X and Y may be dependent. See §4-70 Example 1.

4-31. Worked-out Problems

***Example* 1.** Show that marginal distributions do not determine the joint distributions uniquely.

Solution. Consider an infinite family of bivariate distributions $f_k(x, y)$ defined at the vertices of a unit square.

$$f(0, 0) = (\tfrac{1}{4}) - k = f(1, 1)\,; f(0, 1) = (\tfrac{1}{4}) + k = f(1, 0);\ 0 \le k \le \tfrac{1}{4} \qquad \text{...(1)}$$

Obviously, $P(X = 0) = P(X = 1) = \frac{1}{2}$, $P(Y = 0) = P(Y = 1) = \frac{1}{2}$ ...(2)

Since $P(X = x, Y = y) \ne P(X = x)\, P(Y = y)$. X and Y are dependent. Note that all the infinite bivariate distributions (1) [e.g. $f(0, 0) = f(1, 1) = f(0,1) = f(1, 0) = \frac{1}{4}$; $f(0, 0) = f(1, 1) = 1/8$, $f(0, 1) = f(1, 0) = 3/8$, etc.] possess the same marginals, the uniqueness concept is completely routed.

***Example* 2.** Let X and Y be integer-valued random variables with the joint p.m.f.

$$P(X = m, Y = n) = q^2 p^{m+n-2}\ ;\ n, m = 1, 2, 3 \ldots \qquad (p + q = 1)$$

Are X and Y independent ?

Solution. We use Multi-Stage Rule to get

$$P(X = m) = \sum_{j=1}^{\infty} P(X = m, Y = j) = q^2 p^{m-1} \sum_{j=1}^{\infty} (p)^{j-1} = q^2 p^{m-1} / (1 - p) = q\, p^{m-1}$$

$$P(Y = n) = \sum_{i=1}^{\infty} P(X = i, Y = n) = qp^{n-1}.$$

Since $P(X = m, Y = n) = P(X = m)\ P(Y = n)$, $\forall\ m, n$, we infer that X and Y are independent.

***Example* 3.** Random variables X and Y possess the joint density

$$p(x, y) = \frac{e^{-\lambda}\lambda^x}{y!}\frac{p^y q^{x-y}}{(x-y)!}, p + q = 1,\ y = 0, 1, 2, ..., x;\ x = 1, 2, 3, ...\ \lambda > 0;\ 0 < p < 1.$$

Find the marginal and conditional distributions and evaluate $P(X < 1)$.

Solution. Summing out y, and using binomial expansion, the marginal probability for x is

$$P_X(x) = \sum_{y=0}^{x} p(x, y) = \frac{e^{-\lambda}\lambda^x}{x\,!}\sum_{y=0}^{x}\frac{x!}{y\,!(x-y)\,!}q^{x-y}p^y = \frac{e^{-\lambda}\lambda^x}{x\,!}\sum_{y=0}^{x}\left[\frac{x}{y}\right]q^{x-y}p^y = \frac{e^{-\lambda}\lambda^x}{x\,!}(q+p)^x$$

$$= e^{-\lambda}\lambda^x / x\,!\quad x = 0, 1, 2, ... \qquad ...(1)$$

Summing out x, and using exponential series, the marginal probability for y is

$$P_Y(y) = \sum_{x=0}^{\infty} p(x, y) = \frac{e^{-\lambda}p^y}{y\,!q^y}\sum_{x=y}^{\infty}\frac{(\lambda q)^x}{(x-y)\,!} = \frac{e^{-\lambda}p^y}{y\,!q^y}(\lambda q)^y\sum_{z=0}^{\infty}\frac{(\lambda q)^z}{z\,!}\quad [z = x - y]$$

$$= \frac{e^{-\lambda}(p\lambda)^y}{y!} = e^{\lambda q} = \frac{e^{-\lambda p}(\lambda p)^y}{y!};\ y = 0, 1, 2, ..., x \qquad ...(2)$$

$$P(y \mid x) = \frac{p(x, y)}{P_X(x)} = \frac{e^{-\lambda}\lambda^x}{y!}\frac{p^y\,.q^{x-y}}{(x-y)\,!}\frac{x!}{e^{-\lambda}\lambda^x} = \binom{x}{y}p^y q^{x-y}. \qquad ...(3)$$

$$p(y \mid x) = \frac{p(x, y)}{P_Y(x)} = \frac{e^{-\lambda}\lambda^x}{y!}\frac{p^y\,.q^{x-y}}{(x-y)\,!}\frac{y!}{e^{-\lambda p}(\lambda p)^x} = \frac{e^{-\lambda q}.(\lambda q)^{x-y}}{(x-y)!}. \qquad ...(4)$$

From (1) $P(X < 1) = P(X = 0) = e^{-\lambda}$.

***Example* 4.** The joint p.m.f. of (X, Y) is given by

$$f(x, y) = k\binom{x+y-1}{x}p^x Q^y,\ x \geq 0, y \geq 1,\ 1 > 1 - p > Q > 0.$$

Determine k. Are X, Y independent ?

Solution. Firstly, we determine k and the marginals $f_1(x)$ and $f_2(y)$. Now,

$$\sum_{x=0}^{\infty} f(x, y) = kQ^y\sum_{x=0}^{\infty}\binom{x+y-1}{x}p^x \qquad \left[\text{Recall}:\ (1-T)^{-n} = \sum_{r=0}^{n}\binom{n+r-1}{r}T^r\right]$$

$$= kQ^y(1-p)^{-y} = k(Q/q)^y, \qquad [q = 1 - p,\ P = 1 - Q] \quad ...(i)$$

$$\sum_{x=0}^{\infty}\sum_{y=1}^{\infty} f(x, y) = k\sum_{y=1}^{\infty}\left(\frac{Q}{q}\right)^y = k\,.\frac{Q/q}{1-(Q/q)} \quad \text{(Sum of a G.P.)} \qquad ...(ii)$$

By normality $\Sigma\Sigma f(x, y) = 1$, consequently

$$k = \frac{1-(Q/q)}{(Q/q)} = \frac{P}{Q}\left(1 - \frac{p}{P}\right) \qquad (P + Q = 1,\ p + q = 1) \qquad ...(iii)$$

$$\therefore \quad f_2(y) = \left(1-\frac{Q}{q}\right)\left(\frac{Q}{q}\right)^{y-1}, \; y = 1, 2, 3, \ldots \qquad \text{[by (i) and (iii)]}$$

$$f_1(x) = \sum_y f(x, y) = kp^x . Q \sum_{y=1}^{\infty} \binom{x+y-1}{x} Q^{y-1} \qquad [\text{Put } y-1 = r]$$

$$= kp^x Q . \sum_{r=0}^{\infty} \binom{x+1+r-1}{r} Q^r = kp^x Q . (1-Q)^{-(x+1)} \qquad \text{[Neg-bin Theorem]}$$

$$= \frac{k.Q}{P}\left(\frac{p}{P}\right)^x = \left(1-\frac{p}{P}\right)\left(\frac{p}{P}\right)^x, x = 0, 1, 2, \ldots \qquad \text{[by (iii)]} \qquad \ldots\text{(iv)}$$

We conclude that $X \sim \text{gem}\left(1-\frac{p}{P}\right)$ and $Y \sim \text{gem}\left(1-\frac{Q}{q}\right)$

Since $f(x, y) \neq f_1(x)\, f_2(y)$, we conclude that X and Y are not independent.

Problems with Solutions Provided at the End of the Text

1*. A pair of dice is tossed. Let X be the number on the first die D_1 and S be the sum of the two faces. Determine the distribution (X, S) and show that,

$$P(S = k) = [6 - |7 - k|]/36, \quad k = 2, 3, \ldots, 12.$$

2*. Let X and Y be the minimum and maximum respectively of the results, when two dice are rolled. Find the marginal p.m.f. of Y and $P\{X = x \mid Y = y\}$.

3*. Two cards are drawn at random (without replacement) from a standard 52-card deck. Let X be the number of aces that occur and Y be the number of spades that occur. Derive $p(x, y)$ and compute $P(X > Y)$.

4*. Variates X and Y have the joint probability distribution :

$Y\downarrow$ \ $X\rightarrow$	1	2	3	4	5	6
0	0	0	$2a$	$4a$	$4a$	$6a$
1	$4a$	$4a$	$8a$	$8a$	$8a$	$8a$
2	$2a$	$2a$	a	a	0	$2a$

Obtain the value of a and find the marginal p.m.fs. of X and Y. Evaluate

$$P\{X \leq 1), \quad P(X \leq 1 \mid Y = 2), \quad P(X < 3 \mid Y \leq 4).$$

5*. Independent variates X and Y have distributions :

x	:	1	2	3	4	5
P_X	:	0.1	0.2	0.4	0.2	0.1

y	:	0	1	2	3
P_Y	:	0.2	0.1	0.4	0.3

Find the distribution of $X + Y$, XY, Y/X.

6*. *Partial information.* X and Y are independent variates. Find their distribution when *some* data is provided in the following (L.H.S.) table. [R.H.S. table is solution]

$x \downarrow$ \ $y \rightarrow$	1	2	3	Total P_x
1	0.06	—	—	—
2	—	—	—	0.2
3	—	—	0.3	—
Total P_y	—	0.2	—	—

$x \downarrow$ \ $y \rightarrow$	1	2	3	Total P_x
1	0.06	0.04	0.10	a
2	0.06	0.04	0.10	0.2
3	0.18	0.12	0.30	0.8—a
Total P_y	b	0.2	0.8–b	1

7*. A coin is flipped until heads occur twice. Let X and Y denote, respectively, the trial numbers at which the first and the second heads are observed. If p is the probability of heads occurring at any one trial, find $P\{Y - X = k \mid X = m\}$.

8*. The joint density of (X, Y) is $f(x, y) = (x + 2y)/27$, where x and y can assume only the integer values 0, 1, 2. Find the conditional distribution of Y for $X = x$.

Exercise 4(a)

1. Show why the functions G given below do not represent joint c.d.fs.

   (a) $G(x, y) = 1,\ x + 2y \geq 1;\ G(x, y) = 0,\ x + 2y < 1.$

   (b) $G(x, y) = 1 - e^{-x+y},\ x \geq 0, y \geq 0;\ G(x, y) = 0,$ elsewhere.

   (c) $G(x, y) = 1 - e^{-x-y}, x \geq 0, y \geq 0;\ G(x, y) = 0,$ elsewhere.

   (d) $G(x, y) = 0,\ x < 0, y < 0,$ or $x + y < 1$; $G(x, y) = 1$ otherwise.

   (e) $G(x, y) = 1, x + y \geq 1,\ G(x, y) = 0,\ x + y < 1.$

2. Variates X and Y have the following c.d.f. :

$$F(x, y) = \begin{cases} 0, & x \geq -2 \quad \text{or } y < 0 \\ 0, & -2 \geq x < 3 \text{ and } 0 \leq y < 5 \\ \frac{3}{5}, & x \geq 3 \quad \text{and } 0 \leq y < 5 \\ \frac{2}{5}, & -2 \geq x < 3 \text{ and } y \geq 5 \\ 1, & x \geq 3 \quad \text{and } y \geq 5 \end{cases}$$

   Determine that X and Y indeed have a joint discrete distribution and find the joint probability function.

3. The joint c.d.f. of X and Y is
   $F(x, y) = 1 - e^{-x} + e^{-y} + e^{-x-y}, x \geq 0,\ y \geq 0;\ F(x, y) = 0,$ elsewhere.
   Find (a) $P(-2 < X \leq 3, 1 < Y \leq 2)$, (b) $P(X + Y \geq 3)$, (c) $P(X \geq Y)$, (d) joint p.d.f. of X and Y.

4. Show why the functions f and g defined below cannot represent a joint probability function for any choice of k

$$f(x, y) = \begin{cases} k(2x^2 - xy), & x = 0, 1, 2;\ y = 0, 3 \\ 0, & \text{elsewhere} \end{cases}, \qquad g(x, y) = \begin{cases} kxy, & x, y = -1, 0, 1 \\ 0, & \text{elsewhere} \end{cases}$$

5. X and Y have joint p.d.f.

$$f(x, y) = p^2 q^y, x = 0, 1, 2, ..., y;\ y = 0, 1, 2, ..., 0 < p < 1, p + q = 1.$$

Find the marginal distributions of X and Y and evaluate $P(Y - X \le 1)$.

6. Given the adjoining joint p.d.f., obtain a and marginal distributions of X and Y and conditional distributions of X given $Y = 1$. Evaluate $P(X + Y < 4)$.

$y \downarrow$ \ $x \rightarrow$	1	2	3
1	p	p	$2p$
2	$2p$	$3p$	p

7. X and Y have joint p.d.f.

(a)

$x \downarrow$ \ $y \rightarrow$	−1	1	2	5
−1	a	$3a$	$3a$	a
1	$3a$	$6a$	$3a$	0
5	$4a$	$3a$	0	0

(b)

$x \downarrow$ \ $y \rightarrow$	−1	0	1
0	a	$2a$	a
1	$3a$	$2a$	a
2	$2a$	a	$2a$

For table (a) evaluate : (i) $P(Y$ is even), (ii) $P(XY$ is odd), (iii) $P(X^2Y^2 < 5)$, (iv) $P(Y$ is even and $X^2 < 2)$, (v) $P(0 < X < 5)$.

For table (b) : (i) Obtain marginal distributions of X, Y.

(ii) Obtain conditional distribution of Y give $X = 2$.

8. An urn possesses the distribution $p(1) = 0.2, p(2) = 0.5, p(3) = 0.3, [p(x) = P(X = x)]$.
Find the marginal and conditional distributions for the random experiment consisting of two drawings without replacement from this urn. What is the conditional p.d.f. of X given that $X + Y = 4$?

9. The p.d.f. of X, Y which represent the number of bicycles produced by assembly lines A and B is given below :

$x \downarrow$ \ $y \rightarrow$	0	1	2	3	4	5
0	0	a	$3a$	$6a$	$7a$	$8a$
1	a	$2a$	$4a$	$7a$	$7a$	$9a$
2	a	$2a$	$4a$	$6a$	$6a$	$6a$
3	0	a	$5a$	$5a$	$5a$	$4a$

Find a and the marginal density functions. What is the probability that :

(i) more items are produced by line B than by line A ?

(ii) line B produced 3 items if line A has produced 1 ?

10. A bivariate density $f(x, y)$ is defined for $x = 0, 1, 2, 3, ..., y = 0, 1, 2, 3, ...$ The partial table is given by

$x \downarrow$ \ $y \rightarrow$	0	1	2	3	4	
0	$2a$	$7a$	$8a$	$3a$	$6a$	
1	a	$2a$	$5a$	$8a$	$9a$	
2	a	a	$2a$	$3a$	$4a$	$a = 0.01$
3	$4a$	$2a$	$7a$	$6a$	$8a$	
4	$3a$	$4a$	$7a$	$8a$	a	

Find the marginal density functions and evaluate $P(X + Y \leq 5)$, and $P(X \leq 4 \mid Y \leq 3)$.

11. Two fair dice are tossed simultaneously, X and Y denote the number on the first and second die respectively. Record the sample space for this expeirment. If $U = X - Y$, $D = \mid U \mid$, $S = X + Y$, $T = XY$; find the joint p.d.f. of D and S and of S and T. Also evaluate $P(S = 8)$, $P(S \geq 8)$, $P(U = 0)$, $P(S = 6 \mid Y = 4)$, $P(U = 2)$.

[See p.15 for joint densities. $P(S \geq 8) = 5/12$, $P(S_6 \mid Y = 4) = 1/6$, $P(U = 0) = 1/3$].

12. A number X is chosen at random from among the integers 3, 4, ..., 10. A number Y is then chosen at random from among the even positive integers less than X. Find $P(Y \leq 5 \mid X = x)$ for each x, and $P(Y \leq 5)$.

13. (a) A tosses a skew coin (head probability p) with sides marked 1 and 2. B spins a fair spinner which is evenly graduated from 0 to 3. The pay off of this game is $X = A$'s number – B's number. Find $F_X(x)$.

(b) Select one of the integers 1, 2, 3, 4, 5. After discarding all integers (if any) less than the selected integer, draw one of the remaining integers. Let X and Y denote the numbers obtained on the first and second draws respectively. Construct the joint p.d.f. of X and Y and find $P(X + Y > 7)$.

14. An urn contains 3 red and 2 green balls. A random sample of 2 ball is drawn (a) with replacement, (b) without replacement. Let $X = 0$ if the first ball drawn is green, $X = 1$ otherwise; let $Y = 0$ if the second ball drawn is green, $Y = 1$ otherwise. Find the joint density of (X, Y) and also find the conditional *p.m.f.*'s.

15. The joint p.d.f. of X and Y is : $f(x, y) = (x + y)/21$, $x = 1, 2, 3$; $y = 1, 2$.

Show that the marginal p.d.f's of X and Y are

$$f(x) = (2x + 3)/21,\ x = 1, 2, 3;\ g(y) = (3y + 6)/21,\ y = 1, 2.$$

Show that $P(X \leq 2) = 4/7$ and $P(Y \leq 2) = 1$.

16. The joint p.d.f. of X and Y is $f(x, y) = xy/36$, $x, y = 1, 2, 3$; $f(x, y) = 0$, otherwise.

If $U = X + Y$, $V = X - Y$, find the joint and marginal p.m.f.'s of U and V.

[**Ans.** $f(u, v) = (u^2 - v^2)/144$, over 9 points, $g(2) = p$, $g(3) = 4p$, $g(4) = 10p$, $g(5) = 12p$, $g(6) = 9p$, $p = 1/36$ $h(-2) = 3p$, $h(-1) = 8p$, $h(0) = 14p$, $h(1) = 8p$, $h(2) = 3p$]

17. An electric circuit is closed and opened repeatedly until two electric bulbs B_1 and B_2 fuse and the closures x and y at which (B_1) and (B_2) fuse are recorded. The variation in x and y is described by the density function

$$f(x, y) = \theta^x (1 - \theta)^y;\ x, y = 1, 2, 3, \ldots\ (0 < \theta < 1_q).$$

Find the marginal densities of X and Y. Also find the probability that :

(i) B_1 fuses at the kth closure, (ii) B_1 fuses before B_2, (iii) both bulbs fuse at the same closure.

18. Let $P(X = x, Y = y) = (2^{x-1}3^{y})^{-1}$, $x, y = 1, 2, \ldots$, Show that X and Y are independent.

19. Let $f(x, y) = 2/n(n + 1)$, $x = 1, 2, \ldots, n$, $y = 1, 2, \ldots, x$; $f(x, y) = 0$, otherwise.
Show that X and Y are not independent.

20. Let a and b be constants. Write $P(X = m, Y = n) = f(m, n)$. Define

$$f(m, n) = \frac{e^{-(a+b)} a^m b^{n-m}}{m!\,(n-m)!}, \; m = 0, 1, 2, \ldots; \; n = 0, 1, 2, \ldots; \; f(m, n) = 0 \text{ otherwise.}$$

(i) Show that X and Y are dependent, (ii) Are X and $Y - X$ independent ?

21. X and Y have p.m.f. concentrated at four points

$P(X = 2, Y = 3) = \frac{1}{3}$, $P(X = 2, Y = -1) = a$, $p(X = -1, Y = 3) = b$, $P(X = -1, Y = -1) = \frac{1}{6}$.

If X and Y are independent, show that $a = \frac{1}{3}, b = \frac{1}{6}$ or $a = \frac{1}{6}, b = \frac{1}{3}$.

22. X and Y are random variables with distribution shown in table (a). Are X and Y independent ? Find the missing values.

23. **Partial information :** X and Y are two independent variates with a joint p.d.f. partly shown in table (b). Fill the gaps.

(a)

$X\downarrow$ \ $Y\rightarrow$	0	1	2	
1	0	—	—	0.2
2	—	0.3	—	—
3	0	—	0.1	0.4
	0.1	—	0.2	

(b)

$X\downarrow$ \ $Y\rightarrow$	0	1	
1	1/6	—	
2	—	—	2/3
3	—	—	
		1/2	

24. Suppose X and Y are independent variates with non-negative integer values. Let
$(P(X = n) = a_n, P(Y = n) = b_n, \; n = 0, 1, 2, \ldots$
If $S = X + Y$, show that $P(S = n) = \Sigma a_m b_{n-m}$, $m = 0, 1, 2, \ldots, n$.

25. Let X and Y be discrete variates with $p_{ij} = P(X = x_i, Y = y_j)$. Show that X and Y are indep. iff rank of matrix $(p_{ij}) = 1$.

4.40. *Continuous Random Variable*

Definition. A 2-dimensional random vector (X, Y) is called a *continuous* random vector if there exists a function $f(x, y) \geq 0$ such that for $-\infty < x, y < \infty$, the c.d.f. 'F' of (X, Y) given by

$$F(x, y) = \int_{-\infty}^{x}\left[\int_{-\infty}^{y} f(u, v)\,dv\right]du \qquad \ldots(1)$$

is continuous. The function $f(x, y)$ is called the *joint* p.d.f. of (X, Y).

4-41. *Theorem : Some Properties of Joint Density*

Let $f(x, y) \geq 0$ be the joint p.d.f. of continuous random vector (X, Y) and $F(x, y)$ be the c.d.f. of (X, Y). The following properties hold :

(i) $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x, y)\, dx\, dy = 1.$

(ii) $P(a < X \le b,\ c < Y \le d) = \int_b^a \int_c^d f(x, y)\, dy\, dx$

(iii) $f(x, y) = \partial^2 F(x, y) / \partial x \partial y.$

Proof. (i) The relation between $f(x, y)$ and $F(x, y)$ is

$$F(x, y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f(u, v)\, dv\, du. \qquad ...(1)$$

Since $F(+\infty, +\infty) = 1$, the result (i) flows from (1) trivially.

(ii) $P(a < X \le b,\ c < Y \le d) = F(a, c) - F(b, c) - F(a, d) + F(b, d)$ [Rectangle Rule]

$$\therefore \quad p = \int_{-\infty}^{a}\int_{-\infty}^{c} f\, dy\, dx - \int_{-\infty}^{b}\int_{-\infty}^{c} f\, dy\, dx - \int_{-\infty}^{a}\int_{-\infty}^{d} f\, dy\, dx + \int_{-\infty}^{b}\int_{-\infty}^{d} f\, dy\, dx$$

$$= \int_{-\infty}^{b}\left[\int_{-\infty}^{d} dy - \int_{-\infty}^{c} dy\right] dx - \int_{-\infty}^{a}\left[\int_{-\infty}^{d} f\, dy - \int_{-\infty}^{c} f\, dy\right] dx$$

$$= \int_{-\infty}^{b}\left[\int_{c}^{d} f(x, y)\, dy\right] dx - \int_{-\infty}^{a}\left[\int_{c}^{d} f(x, y)\, dy\right] dx = \int_{a}^{b}\int_{c}^{d} f(x, y)\, dy\, dx.$$

(iii) We apply Fundamental Theorem of Integral Calculus to formula (1) to recover

$$\partial^2 F(x, y) / \partial x\, \partial y = f(x, y). \qquad ...(2)$$

valid at the points of continuity of f. But there are only at most countable number of points where f is not continuous (or F is not differentiable). So we are free to set $f(x, y)$ as per our convenience at these points of discontinuity. Hence (2) holds for every $(x, y) \in R^2$.

4-42. *Individual or Marginal Distributions*

Let (X, Y) be a continuous random vector with joint c.d.f. F and joint p.d.f. 'f'. Then

$$F(x, y) = P(X \le x, Y \le y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f(u, v)\, dv\, du$$

Since $\{X \le x, Y \le y\} = \{X \le x\} \cap \{Y \le y\}$, hence

$$\{X \le x, Y < \infty\} = \{X \le x\} \cap \{Y \le \infty\} = \{X \le x\} \cap \Omega = \{X \le x\}$$

$$\therefore \quad P(X \le x, Y \le \infty\} = P(X \le x) \Rightarrow F_X(x) = F(x, \infty)$$

i.e. $F_X(x) = \int_{-\infty}^{x}\left[\int_{-\infty}^{\infty} f(u, v)\, dv\right] du = \int_{-\infty}^{\infty} g(u)\, du$ where $\int_{-\infty}^{\infty} f(u, v)\, dv = g(u) \ge 0.$

This shows that X is also continuous and p.d.f. of X is given by

$$f_X(x) = g(x) = \int_{-\infty}^{\infty} f(x, y)\, dy \qquad ...(1)$$

Similarly, Y is also a continuous variate and p.d.f. of Y is given by

$$f_Y(y) = \int_{-\infty}^{\infty} f(x, y)\, dx \qquad ...(2)$$

Definition. Let (X, Y) be a 2-dim. continuous random vector with joint p.d.f. $f(x, y)$. Then the individual or *marginal* distributions of X and Y are defined by the p.d.f's.

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y)\, dy; \quad f_Y(y) = \int_{-\infty}^{\infty} f(x, y)\, dx$$

Observation : $P(a \le X \le b) = \int_a^b f_X(x)\, dx = \int_a^b \left(\int_{-\infty}^{\infty} f(x, y)\, dy \right) dx$

4-43. Conditional Distribution Function

The *conditional* c.d.f. of a variate X, given $Y = y$, written $F_{X \mid Y}(x \mid y)$, is defined by

$$F_{X \mid Y}(x \mid y) = \lim_{\epsilon \to 0+} P\{X \le x \mid (y - \epsilon < Y \le y + \epsilon)\} \qquad \ldots(1)$$

provided that the limit in (1) exists.

The *conditional* p.d.f. of X, given $Y = y$, written $f_{X \mid Y}(x \mid y)$, is a non-negative function satisfying

$$F_{X \mid Y}(x \mid y) = \int_{-\infty}^{x} f_{X \mid Y}(t \mid y)\, dt, \ \forall\, x \in R. \qquad \ldots(2)$$

***Theorem* 1.** The conditional p.d.f. $f(x \mid y)$ is a bonafide density function.

Proof. The conditional p.d.f. $f(x \mid y)$ is non-negative by its very definition. Further,

$$\int_{-\infty}^{\infty} f(x \mid y)\, dx = F_{X \mid Y}(\infty \mid y) = 1, \ \text{[by (1)]} \qquad \ldots(3)$$

Since both conditions of a p.d.f. are satisfied, it follows that $f(x \mid y)$ is a bonafide density function.

***Theorem* 2.** The conditional p.d.f. $f(x \mid y)$ is given by

$$f_{X \mid y}(x \mid y) = f(x, y)/f_Y(y)$$

where f_Y is the marginal p.d.f. of Y, $f_Y(y) > 0$, and it is continuous.

Proof. By definition (1), we have

$$F_{X \mid Y}(x \mid y) = \lim_{\varepsilon \to 0+} \frac{P(X \le x, y - \varepsilon < Y \le y + \varepsilon)}{P(y - \varepsilon < Y \le + \varepsilon)} = \lim_{\varepsilon \to 0+} \frac{\int_{-\infty}^{x} \left[\int_{y-\varepsilon}^{y+\varepsilon} f(u, v)\, dv \right] du}{\int_{y-\varepsilon}^{y+\varepsilon} f_2(v)\, dv} \qquad \ldots(i)$$

We observe that

$$\lim_{\varepsilon \to 0+} \frac{1}{2\varepsilon} \int_{y-\varepsilon}^{y+\varepsilon} f(u, v)\, dv = \lim_{\varepsilon \to 0+} \frac{F(u, y + \varepsilon) - F(u, y - \varepsilon)}{2\varepsilon} = F'(u, y)\, f(u, y). \qquad \ldots(iii)$$

We divide the Num. and Den. in (i) by 2ε, proceed to the limits, use (ii) and a similar result for Den. of (i) to get

$$F_{X \mid Y}(x, y) = \frac{\int_{-\infty}^{x} f(u, y)\, du}{f_Y(y)} = \int_{-\infty}^{x} \frac{f(u, y)}{f_Y(y)}\, du.$$

Hence, by definition (2), we get $f(x \mid y) = f(x, y)/f_Y(y)$.

4-44. Worked-out Problems

Terminology. The word i.i.d. stands for *independent and identically distributed.*

***Example* 1.** A probability function which is neither discrete nor continuous.

Solution. Let $f(x, p) = p^x q^{1-x} p^{a-1} q^{b-1} / B(a, b), x = 0, 1,$

where $p + q = 1,\ 0 < p < 1,\ a > 0,\ b > 0.$

This expression is neither a p.m.f. nor p.d.f., because variable x varies only over two discrete values $\{0, 1\}$ and the variable $p \in (0, 1)$. Nevertheless,

$$\sum_{x=0}^{1} f(x,p) = \frac{p^{a-1}q^{b-1}}{B(a,b)} = g(p),\ \text{say},\ \int_0^1 \left(\sum_0^1 f(x,p)\right) dp = \int_0^1 p^{a-1} q^{b-1} dp \,/\, B(a,b) = 1.$$

This shows that $f(x, p)$ is a probability function. Expression for $g(p)$ reveals that $p \sim B_1(a, b)$. The marginal density of X is

$$h(x) = \int_0^1 f(x,p)\,dp = \int_0^1 \frac{p^{a+x-1}(1-p)^{b-x}}{B(a,b)} dp = \frac{B(a+x, b-x+1)}{B(a,b)} = \begin{cases} a/(a+b), & \text{if } x = 1 \\ b/(a+b), & \text{if } x = 0. \end{cases}$$

***Example* 2.** Random variables X, Y have the joint p.d.f.

$f(x, y) = k(x + y),\ 0 < x, y \le 1;\ f(x, y) = 0$, elsewhere.

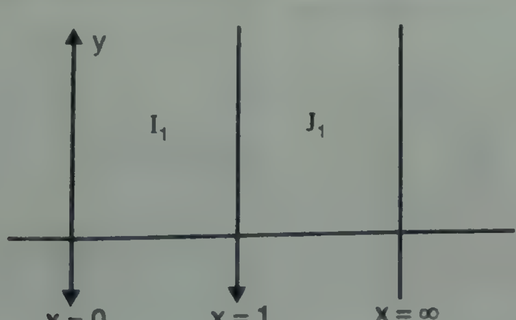
Find k and determine the distribution functions $F(x, y)$, $F(y \mid x)$. Are X, Y independent ? Prove that the marginal densities do not determine the joint density uniquely.

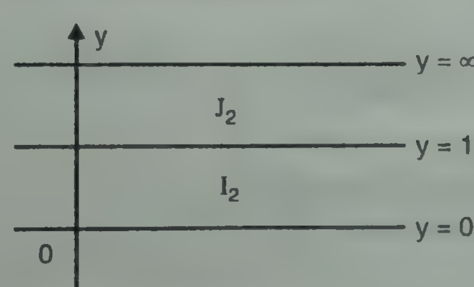
Solution. The norming constant $k > 0$ is obtained by normalization :

$$1 = k \int_0^1 \int_0^1 (x+y)\,dx\,dy = k \int_0^1 [(\tfrac{1}{2}) + y]\,dy = k(\tfrac{1}{2} + \tfrac{1}{2}) = k \,;\text{i.e.}\ k = 1.$$

$$f_1(x) = \int_0^1 (x+y)\,dy = x + \tfrac{1}{2},\quad 0 < x < 1\ ;\quad f_2(y) = y + \tfrac{1}{2}, 0 < y < 1. \qquad \text{...(1)}$$

Since $f_1(x), f_2(y) \ne f(x, y)$, we infer that X and Y are not independent.





While calculating $F(x, y)$, we *drag* indicators $I_1 = I(0 < x < 1)$ and $I_2\ (0 < y < 1)$; $J_1 = I(1 \le x < \infty)$, $J_2 = I\ (1 \le y < \infty)$ along side, a useful trick.

$$F(x,y) = I_1 . I_2 \int_0^y \int_0^x (u+v)\,du\,dv + I_1\, J_2 \int_0^1 \int_0^x (u+v)\,du\,dv + I_2 . J_1 \int_0^y \int_0^1 (u+v)\,du\,dv + 1 . J_1\, J_2$$

$$= \tfrac{1}{2}(x^2 y + x y^2) I_1 I_2 + \tfrac{1}{2}(x^2 + x) I_1\, J_2 + \tfrac{1}{2}(y + y^2) I_2\, J_1 + 1 . J_1 . J_2$$

$$f(y \mid x) = \frac{f(x,y)\, I_1\, I_2}{f_1(x)\, I_1} = \frac{(x+y)\, I_2}{x + (\frac{1}{2})},\qquad 0 < x < 1$$

$$F(y \mid x) = \int_{-\infty}^{y} f_{Y|x}(t \mid x)\, dt = \int_0^y \frac{(x+t)\, dt}{x + (\frac{1}{2})} = \frac{[xy + (y^2/2)]}{x + (\frac{1}{2})}, \qquad 0 < y < 1.$$

Non-uniqueness. Let $g(x, y) = f_1(x)\, I_1 f_2(y)\, I_2 = \left(x + \frac{1}{2}\right)\left(y + \frac{1}{2}\right),\ 0 \le x,\ y \le 1.$

Then $\quad g_1(x) = \left(x + \frac{1}{2}\right) I_1 \int_0^1 \left(y + \frac{1}{2}\right) dy = \left(x + \frac{1}{2}\right) I_1 = f_1(x)$

Similarly, $\quad g_2(y) = \left(y + \frac{1}{2}\right) I_2 = f_2(y).$

Thus, two different continuous distributions f and g possess the **same** marginal distributions f_1 and f_2. It follows that marginal densities do not determine the joint density uniquely.

Example 3. Variates X and Y follow a bivariate *Dirichlet distribution* defined by p.d.f. :

$$f(x, y) = x^{a-1}\, y^{b-1}\, (1 - x - y)^{c-1} / B(a, b, c), \quad x \ge 0,\ y \ge 0,\ x + y \le 1,\ a > 0,\ b > 0,\ c > 0,$$

where $B(a, b, c) = \Gamma(a)\,\Gamma(b)\,\Gamma(c) / \Gamma(a + b + c)$. Find the marginal and conditional densities.

Solution. Obviously : $B(a, b, c) = \dfrac{\Gamma(a)\,\Gamma(b)\,\Gamma(c)}{\Gamma(a+b+c)} = \dfrac{\Gamma(a)\,\Gamma(b+c)}{\Gamma(a+b+c)} \cdot \dfrac{\Gamma(b)\,\Gamma(c)}{\Gamma(b+c)}$

$$= B(a, b + c)\ .\ B(b, c) = B(b, a + c)\ B(a, c).$$

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y)\, dy = \frac{x^{a-1}}{B(a, b+c)} \int_0^{1-x} \frac{y^{b-1}\,(1 - x - y)^{c-1}}{B(b, c)}\, dy \text{ [Put } y = (1-x)\, u,\ dy = (1-x)\, du]$$

$$= \frac{x^{a-1}\,(1-x)^{b+c-1}}{B(a, b+c)} \left[\int_0^1 \frac{u^{b-1}\,(1-u)^{c-1}\, du}{B(b, c)}\right] \qquad \text{[Value of integral is unity]}$$

$$= \frac{x^{a-1}(1-x)^{b+c-1}}{B(a, b+c)}, \qquad 0 \le x \le 1.$$

Similarly, $\quad f_Y(y) = y^{b-1}\,(1 - y)^{a+c-1} / B(b\ .\ a + c),\ 0 \le y \le 1.$

So, if $(X, Y) \sim$ Dir (a, b, c), then $X \sim \beta_1\,(a, b + c)$ and $Y \sim \beta_1\,(b, a + c)$.

Since, $\quad f(x, y) \ne f_1(x)\, f_2(y)$, the variates X and Y are not independent.

$$f_{(y \mid x)} = \frac{f(x, y)}{f_1(x)} = \frac{y^{b-1}\,(1 - x - y)^{c-1}}{B(b, c)\,.\,(1 - x)^{b+c-1}}, \qquad 0 \le y \le 1 - x.$$

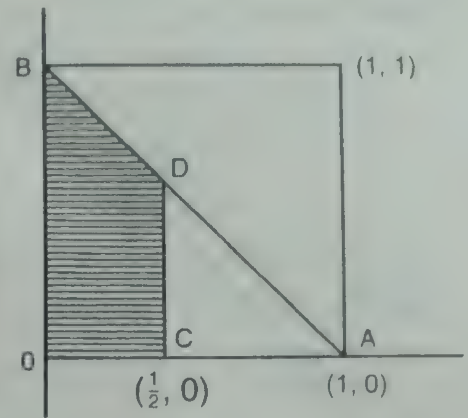
$$f_{(x \mid y)} = \frac{f(x, y)}{f_2(y)} = \frac{x^{a-1}\,(1 - x - y)^{c-1}}{B(a, c)}, \qquad 0 \le x \le 1 - y.$$

Note. $\quad B(a, b, c, d) = \dfrac{\Gamma(a)\,\Gamma(b)\,\Gamma(c)\,\Gamma(d)}{\Gamma(a+b+c+d)}$

$$= B(a, b + c + d)\,(B\,(b, c, d) = B(a, b + c + d)\, B(b + c, d)\, B(b, c).$$

Beta function notation with more than two arguments is due to present authors.

Example 4. The variates X and Y have the joint p.d.f. $f(x, y) = x + y$ for $0 \le x, y \le 1$; $f(x, y) = 0$, elsewhere. Find $P\{2X \le 1 \mid (X + Y) \le 1\}$.



Solution. Here $p = P\{2X \le 1, X + Y \le 1\}/P(X + Y) \le 1)$

$$= P(T)/P(S) \qquad \ldots(1)$$

where $T = \{2x \le 1, x + y \le 1\} = \{0 \le x \le \frac{1}{2}, 0 \le y \le 1 - x\}$ region $OCDB$.

$S = \{x + y \le 1\} = \{0 \le x \le 1, 0 \le y \le 1 - x\}$, triangular region OAB.

$$P(T) = \int_0^{1/2} \int_0^{1-x} (x + y)\, dx\, dy = \frac{1}{2}\int_0^{1/2} (1 - x^2)\, dx = \frac{11}{48}.$$

$$P(S) = \int_0^1 \int_0^{1-x} (x + y)\, dx\, dy = \tfrac{1}{2}\int_0^1 (1 - x^2)\, dx = \tfrac{1}{3}.$$

$$p = \tfrac{11}{16} \qquad \text{[by substitutions into (1)]}$$

Example 5. Random variables X and Y have joint p.d.f.

$$f(x, y) = (2\pi\, xy)^{-1} \exp\left[-\tfrac{1}{4}(\ell n^2 x + \ell n^2 y)\right]; x > 0, y > 0.$$

Are X and Y independent ? Find the marginal density of X.

Solution. We can *factorize* the joint p.d.f as

$$f(x, y) = \left(\sqrt{2}\,\sqrt{2\pi}\, x\right)^{-1} \exp\left[-\tfrac{1}{4}(\ell n\, x)^2\right]\left(\sqrt{2}\,\sqrt{2\pi}\, y\right)^{-1} \exp\left[-\tfrac{1}{4}(\ell n\, y)^2\right]$$

i.e. $\qquad f(x, y) = f_1(x) \,.\, f_2(y); x > 0, y > 0. \qquad \ldots(1)$

We recall that if $Z \sim L\text{-}N\ (\mu, \sigma^2)$, then

$$f_Z(t) = (\sigma\sqrt{2\pi}\, t)^{-1} \exp\left[\frac{-(\ell n\, x - \mu)}{2\sigma^2}\right], \quad t > 0.$$

If follows that $X \sim L\text{-}N\ (0, 2)$ and $Y \sim L\text{-}N\ (0, 2)$ and that X and Y are independent distributed. [by (1)]

Example 6. Let f, g, h be three p.d.f.'s and F, G, H their corresponding c.d.f.'s and suppose a is constant, $\mid a \mid \le 1$. Let

$f_a(x, y, z) = f(x)\, g(y)\, h(z)\, \{1 + a[2F(x) - 1].\ [2G(y) - 1].\ [2H(z) - 1]\}. \qquad \ldots(1)$

Show that f_a is a p.d.f. for each a in $J =]-1, 1[$. Find the marginal p.d.f.'s of the joint density function $f_a(x, y, z)$.

Solution. $0 \le F(x) \le 1 \Rightarrow 0 \le 2F(x) \le 2 \Rightarrow -1 \le 2F(x) - 1 \le 1 \Rightarrow \mid 2F(x) - 1 \mid \le 1$.

Also $\mid a \mid \le 1$, it follows that the product of four numbers in $[-1, 1]$ also lies in $[-1, 1]$.

Thus $\qquad -1 \le a\,(2F - 1)\,(2G - 1)\,(2H - 1) \le 1.$

$\therefore\ 1 + a[2F(x) - 1]\,[2G(y) - 1]\,[2H(z) - 1] \ge 0 \Rightarrow f_a(x, y, z) \ge 0.$

Further, $\displaystyle\int_{-\infty}^{\infty} f(x)\,[2F(x) - 1]\, dx = \tfrac{1}{2}\int_{-1}^{1} t\, dt = 0.\ \ [t = 2F(x) - 1] \qquad \ldots(i)$

Hence integrating (1) between $-\infty < x, y, z < \infty$, we get

$$\iiint f_a(x, y, z)\, dx\, dy\, dz = \int_{-\infty}^{\infty} f(x)\, dx \,.\int_{-\infty}^{\infty} g(y)\, dy \,.\int_{-\infty}^{\infty} h(z)\, dz$$

$$+\int_{-\infty}^{\infty} [2F(x)-1]\, f(x)\, dx \,.\int_{-\infty}^{\infty} [2G(y)-1]\, g(y)\, dy \,.\int_{-\infty}^{\infty} [2H(z)-1]\, h(z)\, dz. \quad \text{...(ii)}$$

On the R.H.S. of (ii) the value of each factor in the first term is unity and the value of each factor in the 2nd term is zero [by (i)]. Hence

$$\iiint f_a(x, y, z)\, dx\, dy\, dz = 1, \quad -\infty < x, y, z < \infty.$$

It follows that f_a is a p.d.f. Integrating (1) w.r.t. y and $z - \infty\ (y, z) < \infty$, we get

$$\iint f_a(x, y, z)\, dy\, dz = f(x)\int g(y)\, dy \int h(z)\, dz + a[2F(x)-1]\, f(x)\int [2G(y)-1]\, g(y)\, dy \int [2H(z)-1]\, h(z)\, dz$$

$$= f(x)\,.\,1.\,1 + a\,[2F(x)-1]\, f(x)\,.\,0\,.\,0 = f(x). \qquad \text{[by (i)]}$$

Thus the x-marginal density of $f_a(x, y, z)$ is $f(x)$. Similarly the marginal p.d.f's of Y and Z are $g(y)$ and $h(z)$ respectively.

Comment. The function $f_a(x, y, z)$ gives an infinite family of joint p.d.f.'s which have the **same** marginal densities. It follows that the knowledge of marginal distributions alone is not sufficient to determine their joint distributions.

Problems with Solutions Provided at the End of the Text

1*. Show that if X, Y, Z are i.i.d. continuous variates, then

$$P(X < Y) = \frac{1}{2!}, \qquad P(X < Y < Z) = \frac{1}{3!}$$

2*. Let X, Y, Z be i.i.d. continuous variates. Evaluate

(i) $P(X > Y \mid X > Z)$, (ii) $P(X > Y \mid X < Z)$, (iii) $P(X > Y \mid Y > Z)$, (iv) $P(X > Y \mid Y < Z)$.

3*. Let $f(x, y) = \lambda^2 e^{-\lambda(x+y)}$, $x \ge 0$, $y \ge 0$, and $f(x, y) = 0$, elsewhere. Find

(a) $P(X < k)$, (b) $P(X > kY)$, (c) $P[(X/Y) \le k]$,

(d) $P[X < Y \mid X < 2Y]$, (e) $P(1 < X + Y < 2)$,

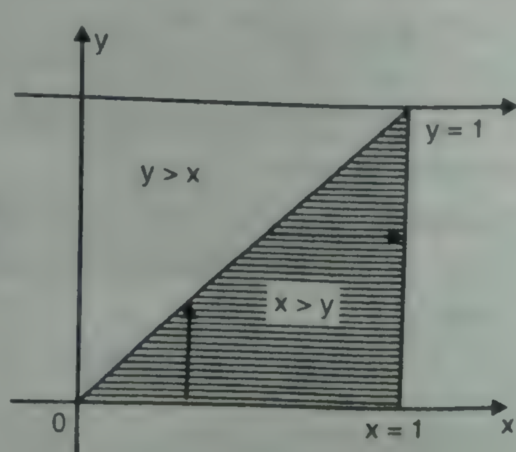
(f) Find m such that when $\lambda = 1$, $P(X + Y < m) = \frac{1}{2}$,

(g) $P[0 < X < 1 \mid Y = 2]$.

4*. The joint p.d.f. of (X, Y) is given by

$f(x, y) = 2,\ 0 < y < x$; $f(x, y) = 0$, elsewhere.

Are X and Y independent ? Find the conditional density functions of Y given $X = x$ and that of X given $Y = y$. Compute $P\{\frac{1}{4} < X < \frac{3}{4} \mid Y = \frac{1}{2}\}$.



5*. The joint p.d.f. of X and Y is given by

$$f(x, y) = kx(x - y),\ 0 < x < 2,\ -x < y < x\ ;\ f(x, y) = 0, \text{ elsewhere.}$$

Evaluate the constant k and the marginal probability density functions of the random variables. Find the also conditional p.d.f. of Y given $X = x$.

6*. The joint p.d.f. of random variables X, Y is

$$f(x, y) = ke^{-(x+y)},\ 0 \le y \le x < \infty\ ;\ f(x, y) = 0, \text{ elsewhere.}$$

Determine k. Find marginal and conditional p.d.f. and evaluate $P(Y \ge 3)$. Are X and Y independent ?

7*. The joint p.d.f. of variates X and Y is

$$f(x, y) = (\tfrac{1}{8})(6 - x - y),\ 0 < x < 2,\ 2 < y < 4\ ;\ f(x, y) = 0, \text{ elsewhere.}$$

Find $P(X < 1, Y < 3)$, $P(X + Y < 3)$, $P(X < 1 \mid Y < 3)$. Find also the marginal and conditional distributions.

8*. Two dimensional r.v. $(X\ Y)$ has the joint density

$$f(x, y) = kxy,\ 0 < x < y < 1\ ;\ f(x, y) = 0, \text{ elsewhere.}$$

Find (i) k and marginal and conditional distributions. Are X and Y independent ?

(ii) Joint distribution function of X and Y. (iii) $P\left\{X < \tfrac{1}{2} \mid Y < \tfrac{1}{4}\right\}$.

9*. Variates X and Y have the joint p.d.f. as :

$$f(x, y) = \frac{1}{8}\pi^2 \sin\frac{1}{2}\pi(x + y),\ 0 < x < 1,\ 0 < y < 1;\ \ f(x, y) = 0, \text{elsewhere.}$$

(a) Find the marginal p.d.f. of X. Are X, Y independent ?

(b) What is the conditional density function $g(y \mid x)$ of Y for the given value x of X ?

(c) Find $P(X > 2Y)$.

10*. If $f_1(x) = qp^{x-1}$, $x = 1, 2, 3, \ldots$, $f_2(y \mid x) = xy^{x-1}$, $0 \le y \le 1$, determine the unconditional density function of Y and the conditional density function of X given Y. Also evaluate $P(Y > \tfrac{1}{2} \mid X = k)$ and $P(X = k \mid Y < \tfrac{1}{2})$.

11*. Let X be the time you get out of bed in the morning (measured in fractions of an hour past 6 A.M.) and let Y be the length of time it takes you to get to your office (in fractions of an hour) after getting up. Assume that the conditional density of Y given $X = x$ and the marginal of X are

$$f_{Y \mid X}(y \mid x) = \frac{2y}{(1 - x)^2}, 0 < x < \frac{2}{3},\ \ 0 < y < 1 - x\ ;\ \ f_X(x) = \left(\frac{81}{26}\right)(1 - x)^2, 0 < x < \frac{2}{3}.$$

Given that it took you 30 minutes to get to your office one morning, find the probability that you got up by 6.15 a.m. that morning. Given that it took you 50 minutes to get there, find the probability that you left your bed later than 6.20 a.m.

Exercise 4(b)

1. (a) If $f(x, y) = k(1 + x + y)^{-3}, x > 0, y > 0$, show that $k = 2$ and $f_1(x) = (1 + x)^{-2}, x > 0$ and $F(x, y) = 1 - (1 + x)^{-1} - (1 + y)^{-1} + (1 + x + y)^{-1}$.

   (b) Determine k such that the joint frequency function of a pair of continuous variables (X, Y) is $f(x, y) = k(xy + 2x + 3y + 6), 0 \le x, y \le 1$; $f(x, y) = 0$, elsewhere.

   Show that X and Y are independent. **[Ans.** $k = 4/35$]

2. The joint distribution of the continuous r.v.s X and Y is
   $f(x, y) = k(x^2 + y^2), 0 \le x \le y \le 1$; $f(x, y) = 0$. elsewhere.
   Determine k and the marginal distributions of X and Y.
   [Ans. $k = 3, f_1(x) = (1 + 3x^2 - 4x^3)\, I(0 < x < 1)$; $f_2(y) = 4y^3\, I(0 < y < 1)$.]

3. The joint p.d.f. of bivariate r.v. (X, Y) is given by
   $f(x, y) = 4(x + y)/5x^3, 0 \le y \le 1, 1 \le x < \infty$; $f(x, y) = 0$, elsewhere.

   (i) Show that $P(0 < y < \frac{1}{2} \mid X = 2) = 9/20$.

   (ii) For what value of a is it true that $P(0 < Y < \frac{1}{2} \mid X > a) = 5/16$? [$a = 1/12$]

4. The joint p.d.f. of bivariate random variable (X, Y) is given by
   $f(x, y) = cy^2, \ 0 < |y| < x, \ 0 < x < 1$; $f(x, y) = 0$, elsewhere.

   (i) Show that $P\{0 < Y < \frac{1}{8} \mid X = \frac{1}{4}\} = \frac{1}{16}$.

   (ii) For what value of a is it true that $P(-a < y < a \mid X = \frac{1}{2}) = \frac{1}{2}$? [$a = (1/2)^{4/3}$]

5. Let $f(x, y) = 1$, when $0 \le x, y \le 1$ and $f(x, y) = 0$, elsewhere. Prove that

   (a) $P\{X < \frac{1}{2}, Y < \frac{1}{2}\} = \frac{1}{4}$, (b) $P\{X > 2Y\} = \frac{1}{4}$, (c) $P\{X^2 + Y^2 < \frac{1}{4}\} = \pi/16$.

   (d) $P\{X > Y$ given $Y > \frac{1}{2}\} = \frac{1}{18}$, (e) $P\{X + Y < 1\} = \frac{1}{2}$, (f) $P(X > \frac{1}{3}) = \frac{2}{3}$,

   (g) $P\{XY > k\} = 1 - k + k \ln k, 0 < k < 1$.

6. The joint p.d.f. of random variables X, Y, is given by
   $f(x, y) = 1/(\pi a^2), x^2 + y^2 \le a^2$; $f(x, y) = 0$, otherwise.

   (i) Show that the marginal distribution of X is $f_1(x) = \left[2\sqrt{a^2 - x^2}/\pi a^2\right] \mathrm{I}(-a \le x \le a)$.

   Find the conditional distribution of Y given $X = x$, where $|x| \le a$. Hence determine if X, Y are mutually independent.

   (ii) Show that $P\{$the distance from the origin of the point selected is not greater than $R\} = R^2/a^2$.

   (iii) Evaluate : $P(|X| < \frac{1}{2} \mid |Y| > \frac{1}{2})$ and $P(X^2 + Y^2 > \frac{3}{4}a^2 \mid X < Y)$.

7. The joint p.d.f. of X and Y is given by
   $f(x, y) = k(y^2 - x^2)\, e^{-y}, -y \le x \le y, 0 < y < \infty$; $f(x, y) = 0$ elsewhere.
   Find k and the marginal densities of X and Y as well as F_X, F_Y, $F(x, y)$.
   [Ans. $k = \frac{1}{8}, f_1(x) = (\frac{1}{4})e^{-|x|}(1 + |x|)\, I(-\infty < x < \infty)$; $f_2(y) = (\frac{1}{6})e^{-y} y^3\, I(0 < y < \infty)$]

8. The joint p.d.f. of X and Y is given by
   $f(x, y) = ky\,(y - x), -y \le x \le y, 0 \le y \le 2$; $f(x, y) = 0$, elsewhere.

(i) Find k, (ii) Find the marginal densities of X and Y.

(ii) The conditional distribution of X, given $Y = y$.

[**Ans.** $k = 1/8; f_1(x) = (1/48)(5x^3 - 12x + 16) I(-2 \le x < 0) + (1/48)(x^3 - 12x + 16) I(0 \le x \le 2)$;
$f_2(y) = (1/4) y^3 I(0 < y < 2)$; $f(x \mid y) = (y - x)/2y^2$; $|x| \le y, 0 < y < 2$].

9. The joint density of X and Y is given by
$f(x, y) = kx^2 (8 - y), x < y < 2x, 0 < x \le 2$; $f(x, y) = 0$, elsewhere.
Find the constant k and the marginal p.d.f.'s of X and Y and the conditional p.d.f.'s

[**Ans.** $k = 5/112$; $f_1(x) = \frac{1}{2} kx^3 (16 - 3x) I(0 < x < 2)$.
$f_2(y) = (7k/24) y^3 (8 - y) I(0 \le y \le 2) + (k/24)(64 - y^3)(8 - y) I(2 \le y \le 4)$]

10. The joint p.d.f. of X and Y is given by
$f(x, y) = 4xy, 0 < x < 1, 0 < y < 1$; $f(x, y) = 0$, elsewhere.
Show that X and Y are independent. Also prove that the probability that exactly one of the random variables is greater than a given constant k, where $k < 1$ is $2k^2 (1 - k^2)$.
Prove further that $F_X(t) = F_Y(t) = t^2$, $F(x, y) = x^2 y^2$.

11. (a) Let the joint p.d.f. of X, Y be $f(x, y) = \frac{1}{2}$, $|x \pm y| \le 1$.

Show that X and Y are identically distributed : $f(t) = 1 - |t|, -1 \le t \le 1$ but are not independent.

(b) If X and Y are i.i.d variates with p.d.f. $f(t)$, that is symmetric about $t = 0$, show that
$p = P\{|X + Y| \le 2|X|\} > \frac{1}{2}$.

Evaluate p for some different symmetric density $f(z)$.

12. The joint p.d.f. of X and Y is $f(x, y) = kxe^{-x(y+1)}, x \ge 0, \ y \ge 0$; $f(x, y) = 0$, elsewhere.
Find (a) the constant k, (b) the marginal densities of X and Y, (c) the conditional densities.

[**Ans.** $k = 1, f_1(x) = e^{-x} I(x \ge 0)$; $f_2(y) = (1 + y)^{-2} I(y \ge 0)$; $f(y \mid x) = xe^{-xy} I(y \ge 0)$]

13. The joint p.d.f. of X and Y is $f(x, y) = k(xy + e^x), 0 \le x, y \le 1$; $f(x, y) = 0$, elsewhere.
Find (i) k, (ii) the marginal densities of X and Y, (iii) the conditional densities, (iv) Are X and Y independent ? [**Ans.** $k = 4(4e - 3)^{-1}$; $f_1(x) = (k/2)[x + 2e^x], f_2(y) = (k/2)(y + 2e - 2)$]

14. The joint p.d.f. of X and Y is $f(x, y) = k(x + y), 0 \le x \le a, 0 \le y \le b$; $f(x, y) = 0$ elsewhere.
Find (i) k, (ii) the marginal and conditional densities of X and Y, (iii) $P(X > a \mid Y > b/2)$.

[**Ans.** $k = 2/ab(a + b)$; $f_1(x) = (2x + b)/a(a + b)$; $f(x \mid y) = (2x + y)/a(a + 2y)$; $p = 0$]

15. If $f(x, y) = 3x^2y + 3xy^2, 0 \le x \le 1, 0 \le y \le 1$; zero elsewhere.

(i) Find the marginal and conditional densities.

(ii) Evaluate $P(\frac{1}{2} \le X \le \frac{3}{4} \mid \frac{1}{3} \le Y \le \frac{2}{3})$.

[**Ans.** $f_1(t) = f_2(t) = \frac{1}{2} t(3t + 2)$; $p = 311/(32)^2, f_1(x) f_2(y) \ne f(x, y)$]

16. (a) $f(x, y) = 24y(1 - x), 0 \le x \le 1, 0 \le y \le x$; $f(x, y) = 0$ elsewhere.

(b) $f(x, y) = 2xy, 0 \le x, y \le 1, x \ge y$; $f(x, y) = 6xy, 0 \le x, y \le 1, x < y$.

(i) Find the marginal and conditional densities. Are X and Y independent ?

(ii) For (b), evaluate $P(A \cup B)$, where $A = \{X \le \frac{1}{2}\}, B = \{Y \le \frac{1}{2}\}$.

[**Ans.** (a) $f_1(x) = 12x^2 (1 - x) I(0 < x < 1), f_2(y) = 12y(1 - y)^2 I(0 < y < 1)$
$f(x + y) = 2(1 - x)(1 - y)^{-2} I(y < x < 1)$; $f(y \mid x) = 2yx^{-2} I(0 < y < x)$.
(b) $f_1(x) = x(3 - 2x^2) I(0 < x < 1), f_2(y) = (y + 2y^3) I(0 < y < 1)$, $p = 7/16$.]

17. The joint p.d.f. of X and Y is $f(x, y) = (3/2)y^2$, $0 \le x \le 2$, $0 \le y \le 1$; $f(x, y) = 0$, elsewhere.
    (i) Find the marginal and conditional distributions. Are X and Y independent ?
    (ii) Find $P(X \le Y)$, $P(2Y \le X)$ [**Ans.** $f_1(x) = \frac{1}{2}$, $f_2(y) = 3y^2$, Indep. $P(Y \le kX) = 2k^3$]

18. The conditional and marginal densities are

$$f(y \mid x) = \frac{3x + y}{3x + 1} e^{-y}, \ y > 0, \quad x > 0; \quad f_X(x) = \frac{3x + 1}{4} e^{-x}, \quad x > 0.$$

Find the joint density of X and Y and the conditional distribution of X given $Y = y$.
[**Ans.** $f(x, y) = (1/4)(3x + y) e^{-(x + y)} u(x) u(y)$; $f(x \mid y) = (3x + y)(3 + y)^{-1} e^{-x}$, $x > 0$].

19. Let $f(x \mid y) = y^x e^{-y}/x!$, $x = 0, 1, 2, \ldots$; $y > 0$; $f(x \mid y) = 0$, elsewhere.
    (a) Show that $f(x \mid y)$ is the conditional p.d.f. of X given $Y = y$,
    (b) Let the marginal p.d.f. of Y be $f_2(y) = \lambda e^{-\lambda y}$, $y > 0$; $f_2(y) = 0$, elsewhere ($\lambda > 0$).
    Find the joint p.d.f. of X and Y and the marginal p.d.f. of X.
    [**Ans.** (a) $f(x \mid y) \sim$ pois (y). $f(x, y) = \lambda e^{-(1+\lambda)y} y^x/x!$. $f_1(x) = pq^x$, $p = \lambda/(1 + \lambda)$]

20. If $f(x, y) = kx^{m-1}(y - x)^{n-1} e^{-y}$ for $0 < x < y < \infty$, show that $X \sim$ gam $(1, m)$ and $Y \sim$ gam $(1, m + n)$.

21. Let $f_X(x) = 2^{-x}$, $x = 1, 2, \ldots$ and $f(y \mid x) = x(1 - y)^{x-1}$, $0 \le y \le 1$. Show that the unconditional distribution of Y is $g(y) = 2(1 + y)^{-2}$, $0 \le y \le 1$.

22. The time taken by a garage to repair a car is a continuous random variable with p.d.f.

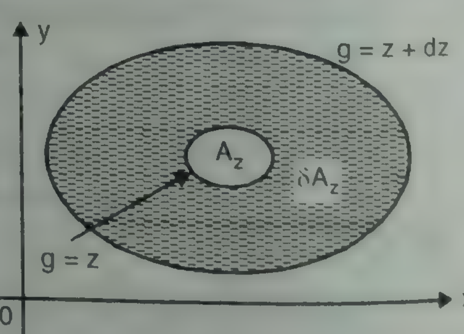
$$f(x) = \tfrac{3}{4} x(2 - x), \quad 0 \le x \le 2; \ f(x) = 0, \text{ elsewhere.}$$

If, on leaving his car, a motorist goes to keep an engagement lasting for a time Y, where Y is a continuous variate, independent of X, with p.d.f.

$$g(y) = y/2, \quad 0 \le y \le 2\ ; \ g(y) = 0, \text{ elsewhere,}$$

show that the probability that the car will not be ready on his returns is 3/10.

4-50. *One Function of Two Random Variables. Standard Formulas*



Let the r.v. Z be a function of two r.v.s. X and Y, say $Z = g(X, Y)$ and suppose $F_Z(z)$ and $f_Z(z)$ are the c.d.f. and p.d.f. of Z respectively. We want to determine these functions in terms of $g(x, y)$ and the joint p.d.f. $f(x, y)$ of X and Y.

Distribution Function. Let A_Z be the region of the x-y plane such that $g(x, y) \le z$. This region can be multipli-connected. Since $\{Z \le z\} = \{(X, Y) \in A_z\}$, it follows that $P(Z \le z)$ consisting of all outcomes ω such that $Z(\omega) = z$ equals the probability mass in the region A_Z. Hence

$$F_z(z) = P\{Z \le z\} = \{(X, Y) \in A_z\} = \iint_{A_z} f_z(x, y)\, dx\, dy \qquad \ldots(1)$$

Probability Density Function. The density $f_z(z)$ can be found by differentiating $F_z(z)$. It may be directly found by determining the region δA_z of the x-y plane such that $z < g(x, y) \le z + dz$, since

$$\{z < Z \le z + dz\} = \{(X, Y) \in \delta A_z\} \qquad \ldots(i)$$

But $$P(z < Z \le z + dz\} = f_z(z)\, dz \quad [\text{by Def.}] \qquad \ldots(ii)$$

As we see from (i), the probability in (ii) equals the masses in the region δA_z. Hence

$$f_z(z) = P(z < Z \le z + az) = \iint_{\delta A_z} f_{X,Z}(x, y)\, dx\, dy \qquad ...(2)$$

Some illustrations of these results follow.

4-51. Distribution of Sum, Difference, Product and Quotient of Two Variates

(i) Distribution of Sum : $Z = X + Y$

Let $Z = X + Y$, and set $A_z = \{(x, y) : x + y \le z\}$: the shaded region in the figure

$$F_z(z) = \iint_{A_z} f(x, y)\, dx\, dy = \int_{-\infty}^{\infty} \left[\int_{-\infty}^{z-x} f(x, y)\, dy\right] dx \qquad ...(i)$$

Put $x + y = u$, $dy = du$ and rewrite (i) after interchanging the order of integration :

$$F_z(z) = \int_{-\infty}^{\infty} \left(\int_{-\infty}^{z} f(x, u - x)\, du\right) dx = \int_{-\infty}^{z} \left(\int_{-\infty}^{\infty} f(x, u - x)\, dx\right) du \qquad ...(ii)$$

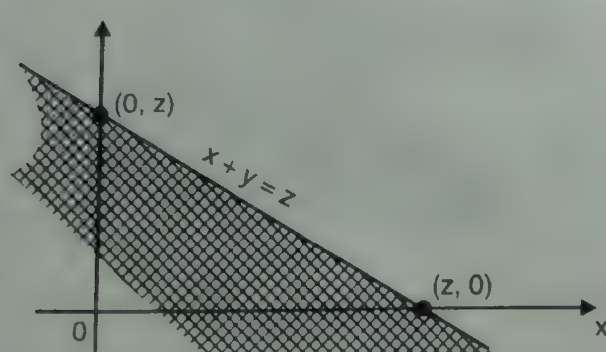
Differentiating this Integral w.r.t. 'z' we obtain the p.d.f. of Z as

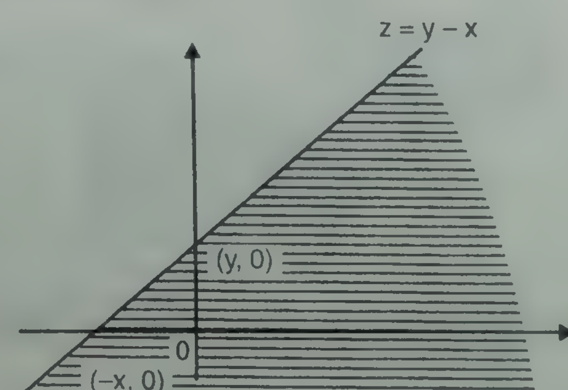
$$f_z(z) = \int_{-\infty}^{\infty} f_{X,Y}(x, z - x)\, dx \qquad ...(1)$$

If X and Y are independent, then (1) gives

$$f_z(z) = \int_{-\infty}^{\infty} f_1(x)\, f_2(z - x)\, dx. \qquad ...(2)$$

The p.d.f. of Z given by (2) is called *convolution* of f_X and f_Y.





(ii) Distribution of Difference : $Z = Y - X$

Let $Z = Y - X$, and set $B_z = \{(x, y) : y - x \le z\}$, the shaded region in the figure. Now

$$F_Z(z) = \iint_{B_z} f(x, y)\, dx\, dy = \int_{-\infty}^{\infty} \left[\int_{-\infty}^{z+x} f(x, y)\, dy\right] dx \qquad ...(i)$$

Put $y = u + x$, $dy = du$ and rewrite (i), after changing the order of integration :

$$F_Z(z) = \int_{-\infty}^{\infty} \left[\int_{-\infty}^{Z} f(x, u + x)\, du\right] dx = \int_{-\infty}^{Z} \left[\int_{-\infty}^{\infty} f(x, x + u)\, dx\right] du \qquad ...(ii)$$

Differentiating this integral w.r.t. z, we obtain the p.d.f. of Z as

$$f_Z(z) = \int_{-\infty}^{\infty} f(x, x + z)\, dx, \quad -\infty < z < \infty \qquad ...(1)$$

If X and Y are independent, this gives

$$f_Z(z) = \int_{-\infty}^{\infty} f_1(x).f_2(x+z)\,dx, \quad -\infty < z < \infty \qquad \text{...(2)}$$

(iii) *Distribution of Product* : Z = *XY*

Here : $\{Z \le z\} = \{XY \le z\} = \{(x, y) : x < 0, y > z/x\} \uplus \{(x, y) : x > 0, y < z/x\}$

$$\therefore \quad F_Z(z) = \int_{-\infty}^{0}\left(\int_{z/x}^{\infty} f(x, y)\,dy\right)dx + \int_0^{\infty}\left(\int_{-\infty}^{z/x} f(x, y)\,dy\right)dx. \qquad \text{...(i)}$$

Put $y = u/x$, $dy = du/x$ and rewrite (i) after changing the order of integration :

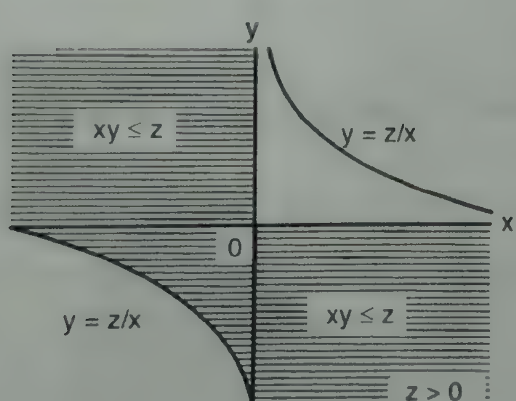
$$\therefore \quad F_Z(z) = \int_{-\infty}^{0}\left(\int_{z}^{\infty} f\left(x, \frac{u}{x}\right)\frac{du}{x}\right)dx + \int_0^{\infty}\left(\int_{-\infty}^{z} f\left(x, \frac{u}{x}\right)\frac{du}{x}\right)dx$$

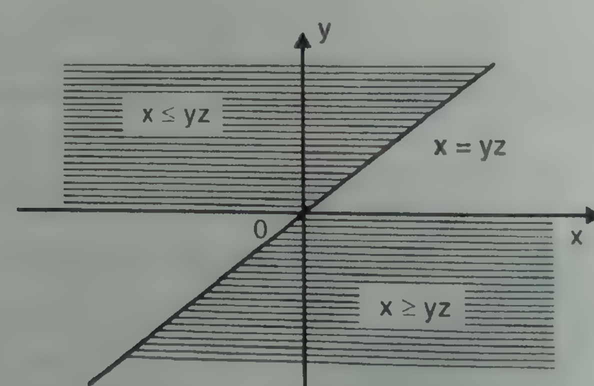
$$= \int_{z}^{\infty}\left(\int_{-\infty}^{0} f\left(x, \frac{u}{\lambda}\right)\frac{dx}{x}\right)du + \int_{-\infty}^{z}\left(\int_0^{\infty} f\left(x, \frac{u}{x}\right)\frac{dx}{x}\right)du \qquad \text{...(ii)}$$

Differentiating this Integral w.r. ... 'z', we obtain the p.d.f. of Z as

$$f_z(z) = \int_{-\infty}^{0} f\left(x, \frac{z}{x}\right)\frac{dx}{(-x)} + \int_0^{\infty} f\left(x, \frac{z}{x}\right)\frac{dx}{x}$$

$$= \int_{-\infty}^{\infty} f_{X,Y}\left(x, \frac{z}{x}\right)\frac{dx}{|x|} \qquad \text{...(1)}$$





If X and Y are independent, then (1) provides

$$f_z(z) = \int_{-\infty}^{\infty} f_1(x)\, f_2\left(\frac{z}{x}\right)\frac{dx}{|x|}. \qquad \text{...(2)}$$

(iv) *Distribution of Quotient* : Z = *X/Y*

Here : $\{Z \le z\} = \{X/Y \le z\} = (x, y) : y > 0, x \le yz\} \uplus \{(x, y) : y < 0, x \ge yz\}$

$$\therefore \quad F_Z(z) = \int_0^{\infty}\left(\int_{-\infty}^{yz} f(x, y)\,dx\right)dy + \int_{-\infty}^{\infty}\left(\int_{yz}^{\infty} f(x, y)\,dx\right)dy \qquad \text{...(i)}$$

Put $x = uy$, $dx = y\,du$ and rewrite (i) after changing the order of integration :

$$F_Z(z) = \int_0^{\infty}\left(\int_{-\infty}^{z} f(uy, y)\,y\,du\right)dy + \int_{-\infty}^{0}\left(\int_{z}^{\infty} f(uy, y)\,y\,du\right)dy$$

$$= \int_{-\infty}^{z}\left(\int_0^{\infty} f(uy, y)\,y\,dy\right)du + \int_{z}^{\infty}\left(\int_{-\infty}^{0} f(uy, y)\,y\,dy\right)du \qquad \text{...(ii)}$$

Differentiating this Integral w.r.t. 'z' we obtain the p.d.f. of Z as

$$f_Z(z) = \int_0^{\infty} f(yz, y)\, y\, dy + \int_{-\infty}^{0} f(yz, y)(-y)\, dy = \int_{-\infty}^{\infty} f_{X,Y}(yz, y)\,|y|\, dy. \quad \ldots(1)$$

If X and Y are independent variates, (1) reduces to

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(yz)\, f_Y(y)\,|y|\, dy. \quad \ldots(2)$$

4-52. *Convolution of Two Distribution Functions*

Definition. A function f on $\mathbb{R}$ is said to be convolution of two d.f.'s F_1 and F_2, written $F = F_1 * F_2$, if

$$F(x) = \int_{-\infty}^{\infty} F_1(x-y)\, dF_2(y), \quad x \in R. \quad \ldots(1)$$

Theorem. The convolution function F of two d.f.'s F_1 and F_2 is itself a distribution function.

Proof.

$$F(x) = \int_{-\infty}^{\infty} F_1(x-y)\, dF_2(y), \quad x \in R \quad \text{(Def.) (1)}$$

As F_1 is increasing, so F must increase. Since the function F_1 is bonded by 1, $[0 \le F_1(t) \le 1]$ and $\lim_{t\to\infty} F_1(t) = F_1(\infty) = 1$,

$$\therefore \quad \lim_{x\to\infty} F(x) = \int_{-\infty}^{\infty} \lim_{x\to\infty} F_1(x-y)\, dF_2(y) = \int_{-\infty}^{\infty} dF_2(y) = 1$$

i.e. $F(\infty) = 1$. Similarly, $F(-\infty) = 0$.

The function F_1, being a c.d.f. is right continuous i.e. $\lim_{t\to a^+} F_1(t) = F_1(a)$, $[0 \le F_1(t) \le 1]$

$$\therefore \quad \lim_{x\to a^+} F(x) = \lim_{x\to a^+} \int_{-\infty}^{\infty} F_1(x-y)\, dF_2(y) = \int_{-\infty}^{\infty} \lim_{x\to a^+} F_1(x-y)\, dF_2(y)$$

$$= \int_{-\infty}^{\infty} \lim_{x\to a^+} F_1(a-y)\, dF_2(y) = F(a), \quad \text{[by (1)]}.$$

This shows that F is right continuous.

Since F satisfies all the properties of c.d.f., it is necessarily a distribution function.

Remark. In the above proof, we have used the fact that if the functions involved are bounded, then the integral (Reimann-Stieltjes) and the 'limit' can be interchanged. This fact is called "dominated convergence theorem".

4-53. *Distribution of the Maximum and the Minimum of Two Variates*

(i) Let $U = \min(X, Y)$; then

$F_U(u) = P\{U \le u\} = 1 - P\{U > u\} = 1 - P\{\min(X, Y) > u\} = 1 - P\{X > u, Y > u\}$. ...(1)

Of course, $-\infty < u < \infty$. If X and Y independent, then

$P\{X > u, Y > u\} = P(X > u)\, P(Y > u) = [1 - P(X \le u)]\,[1 - P(Y \le u)] = [1 - F_X(u)]\,[1 - F_Y(u)]$

$$\therefore \quad F_U(u) = F_X(u) + F_Y(u) - F_X(u)\, F_Y(u) \quad \ldots(2)$$

So $f_u(u) = f_X(u) + f_Y(u) - F_X(u) - f_Y(u)\, f_X(u)\, F_Y(u).$ $\quad [f(t) = F'(t)]$ ...(3)

(ii) Let $V = \max(X, Y)$; then

$$F_V(v) = P\{V \le v\} = P\{\max(X, Y) \le v\} = P\{X \le v, Y \le v\} = F_{X,Y}(v, v) \quad ...(1)$$

If X and Y are independent, then

$$F_V(v) = F_X(v) \,.\, F_Y(v) \qquad -\infty < v < \infty \quad ...(2)$$

$$= F_X(v) \,.\, f_Y(v) + f_X(v)\, F_Y(v). \quad [f(t) = F'(t)] \quad ...(3)$$

4-54. *Distribution of Order Statistics*

Let $X_1, X_2, ..., X_n$ be i.i.d. continuous variates with common c.d.f. and p.d.f. as F and f so that $f = F'$. Consider a linear transformation that places $(X_1, ..., X_n)$ in non-decreasing order. Thus

$$Y_1 \equiv X_{(1)} = \text{smallest of } \{X_1, ..., X_n\}$$
$$Y_2 \equiv X_{(2)} = \text{second smallest of } \{X_1, ..., X_n\}$$
$$.....$$
$$Y_k \equiv X_{(k)} = k\text{th smallest of } \{X_1, ..., X_n\}$$
$$Y_n \equiv X_{(n)} = \text{largest of } \{X_1, ..., X_n\}$$

Then $Y_j = X_{(j)}$, $j = 1, 2, ..., n$ are known as the *order statistics* of $\{X_1, ..., X_n\}$.

Distribution of $X_{(r)}$:

$$P\{X_{(r)} < x\} = P\{\text{At least } r \text{ out of } n \text{ variates are less than } x\} \quad ...(1)$$

Write temporarily : $\quad p = P(X \le x) = F(x),\ q = 1 - p = 1 - F(x),\ p' = F'(x) = f(x)$
$q' = -f'(x)$. Then (1) can be written as $[Y_r = X_{(r)}]$

$$P\{Y_r < x\} = \sum_{k=r}^{n} \binom{n}{k} q^{n-k} p^k \qquad [p = \text{success's probability}]$$

Differentiating w.r.to x and taking $p' = f(x)$ common out of sums, we get

$$g_r(y_r) = f(x)\left\{\sum_{k=r}^{n} \frac{n!\, p^{k-1} q^{n-k}}{(n-k)!\,(k-1)!} - \sum_{k=r}^{n-1} \frac{n!\, p^k q^{n-k-1}}{k!\,(n-k-1)!}\right\}$$

In the second term, put $k = j - 1$ to get

$$g_r(y_r) = f(x)\left\{\sum_{k=r}^{n} \frac{n!\, p^{k-1} q^{n-k}}{(n-k)!\,(k-1)!} - \sum_{j=r+1}^{n} \frac{n!\, p^{j-1} q^{n-j}}{(n-j)!\,(j-1)!}\right\}$$

$$= f(x)\left\{\sum_{k=r}^{n} T_k - \sum_{k=r+1}^{n} T_k\right\}, \qquad \left[T_k = \frac{n!\, p^{k-1} q^{n-k}}{(n-k)!\,(k-1)!},\ \text{dummy } j = k\right]$$

$$= f(x) \,.\, T_r$$

$$= \frac{n!}{(n-r)!\,(r-1)!} f(x) \,.\, p^{r-1} q^{n-r}$$

$$= \binom{n}{r-1, 1, n-r} [F(x)]^{r-1} \,.\, [f(x)] \,.\, [1 - F(x)]^{n-r}$$

We can rewrite this result as

$$\boxed{g_r(y_r) = \binom{n}{r-1,1,n-r}\left[\int_{-\infty}^{y_r} f(x)\,dx\right]^{r-1}[f(y_r)].\left[\int_{y_r}^{\infty} f(x)\,dx\right]^{n-r}.}$$

Special Cases :

$$Y_1 = \min\{X_1, ..., X_n\},\ Y_n = \max\{X_1, ..., X_n\},\ \tilde{X} = \text{median}\,(X_1, ..., X_n)$$

$$g_1(y_1) = nf(y_1).\left[\int_{y_1}^{\infty} f(x)\right]^{n-1}, -\infty < y_1 < \infty$$

$$g_n(y_n) = nf(y_n)\left[\int_{-\infty}^{y_n} f(x)\,dx\right]^{n-1}, -\infty < y_n < \infty$$

Suppose $n = 2m + 1$; then $\tilde{X} = y_{m+1}$

$$\therefore \qquad h(\tilde{x}) = \frac{(2m+1)!}{m!\,m!}\left[\int_{-\infty}^{\tilde{x}} f(x)\,dx\right]f(\tilde{x})\left[\int_{\tilde{x}}^{\infty} f(x)\,dx\right]^{m}, \ -\infty < \tilde{x} < \infty$$

Note that for $n = 2m$, $\tilde{X} = \frac{1}{2}(y_m + y_{m+1})$

4-55. *Worked-out Problems*

***Example* 1.** The joint p.d.f. of X and Y is

$$f(x, y) = \lambda^2 e^{-\lambda(x+y)},\ \lambda > 0,\ x > 0,\ y > 0;\ f(x, y) = 0, \text{ elsewhere.}$$

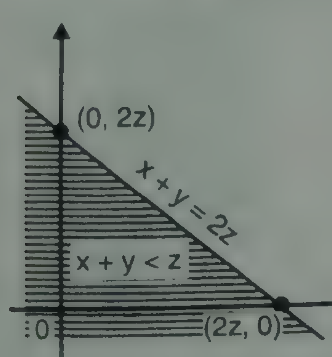
Show that the p.d.f. of $Z = \frac{1}{2}(X + Y)$ is $g(z) = 4\lambda^2 ze^{-2\lambda z}$, $0 < z < \infty$.

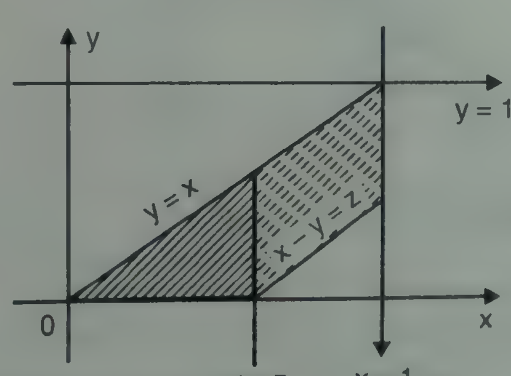
Solution. $G(z) = P(Z \le z) = P(X + Y \le 2z) = P(0 < X \le 2z, 0 \le Y \le 2z - x)$ [(X, Y) indep.]

$$= \int_0^{2z}\int_0^{2z-x} \lambda^2 e^{-(x+y)}dx\,dy = \int_0^{2z} \lambda e^{-\lambda x}[1 - e^{-\lambda(2z-z)}]dx = \lambda\int_0^{2z}(e^{-\lambda x} - e^{-2\lambda z})dx = 1 - (1 + 2\lambda z)e^{-2\lambda z}$$

$$\therefore \qquad g(z) = G'(z) = 4\lambda^2 ze^{-2\lambda z}.$$

Comments. The mgf method dissolves the issue most quickly.





***Example* 2.** The joint p.d.f. of X and Y is given by

$$f(x, y) = 3x,\ 0 < y < x,\ 0 < x < 1;\ f(x, y) = 0, \text{ elsewhere.}$$

Find the p.d.f. of $Z = X - Y$.

Solution.

$$G(z) = P(Z \le z) = P(X - Y \le z) = P(Y \ge X - z) = \int_0^1 \int_{x-z}^{x} f(x, y)\, dx\, dy$$

$$= \int_0^z 3x\, dx \int_0^x dy + \int_z^1 3x\, dx \int_{x-z}^{x} dy$$

$$= 3\int_0^z x^2\, dx + 3z\int_z^1 x\, dx = (3z - z^3)/2$$

$\therefore$

$$g(z) = G'(z) = \tfrac{3}{2}(1 - z^2), 0 < z < 1.$$

Note. Examples 1 and 2 can be dealt by Jacobian method rather quickly.

Example 3. Let X, Y, Z be i.i.d. $N(0, 1)$ variates. Find the p.d.f. of $W = \sqrt{X^2 + Y^2 + Z^2}$.

Solution. Here $f_X(x) = (\sqrt{2\pi})^{-1} e^{-x^2/2}$, $-\infty < x < \infty$, etc. Now the joint c.d.f. differential. of X, Y, Z is

$$dF(x, y, z) = f_1(x)\, dx\, f_2(y)\, dy\, f_3(z)\, dz \qquad [f(x, y, z) = f_1(x) f_2(y) f_3(z)]$$

$$= (1/\sqrt{2\pi})^3 e^{-(x^2+y^2+z^2)/2} dx\, dy\, dz. \qquad \text{...(i)}$$

This form suggests using spherical coordinates (r, ϕ, θ) given by $x = r \sin\phi \cos\theta$, $y = r \sin\phi \sin\theta$, $z = r\cos\phi$.

Here $\partial(x, y, z)/\partial(r, \phi, \theta) = r^2 \sin\phi$, $dx\, dy\, dz = r^2 \sin\phi\, dr\, d\phi\, d\theta$

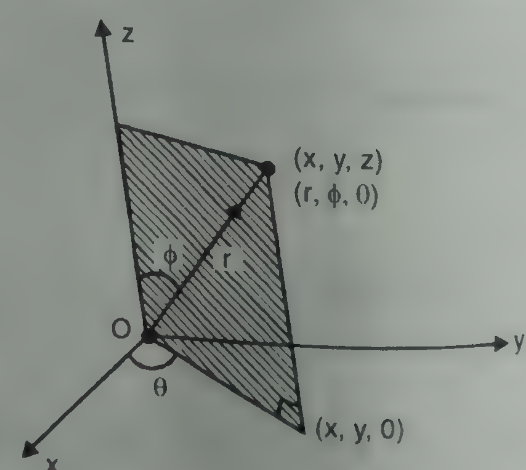
$$x^2 + y^2 + z^2 = r^2,\ 0 \le \phi \le \pi, 0 \le \theta \le 2\pi, 0 \le r \le w.$$

Now $\quad F_W(w) = P\{W \le w\} = P\{X^2 + Y^2 + Z^2 \le w^2\}$

$$= \iiint_{R_W} f(x, y, z)\, dx\, dy\, dz, \quad R_W = \{x, y, z : x^2 + y^2 + z^2 \le w^2\}$$

$$= \int_0^{2\pi}\int_0^{\pi}\int_0^{w} \left(\frac{1}{\sqrt{2\pi}}\right)^3 e^{-r^2}.(r^2 \sin\phi)\, dr\, d\theta\, d\phi, \qquad \text{[by (i) in sphericals]}$$

$$= \sqrt{\frac{2}{\pi}} \int_0^w r^2 e^{-r^2} dr$$



Differentiating this w.r.to w gives

$$f_W(w) = \frac{d}{dw} F_W(w) = \begin{cases} \sqrt{\dfrac{2}{\pi}} w^2 e^{-w^2}, & w > 0 \\ 0, & \text{if } w < 0 \end{cases}$$

$[W \sim \chi_{(3)}]$.

Problems with Solutions Provided at the End of the Text

1*. The r.v.'s X and Y are independent with $Y \sim U(0, 1)$. If $Z = X + Y$, prove that

$$F_Z(z) = F_1(z) - F_1(z - 1).$$

2*. If X and Y are i.i.d Expo (λ) variates, find the p.d.f. of (a) $U = \min(X, Y)$, (b) $V = \max(X, Y)$.

4-60. *Two Functions of Two Random Variables*

Let the variates U and V be functions of random variables X and Y say, $U = g(X, Y)$, $V = h(X, Y)$ and suppose $F_{U,V}(u, v)$ and $f_{U,V}(u, v)$ are the joint c.d.f. and p.d.f. of U and V respectively. We want to determine these functions in terms of $g(x, y)$ and $h(x, y)$ and the joint p.d.f. $f_{X,Y}(x, y)$ of X and Y.

Joint Distribution Function. Let B_{UV} be the region of x-y plane such that $g(x, y) \le u$ and $h(x, y) \le v$. Since $\{U \le u, V \le v\} = \{(X, Y) \in B_{UV}\}$ it follows that $P(U \le u, V \le v)$ consisting of all outcomes ω such that $U(\omega) \le u$, $V(\omega) \le v$, equals the probability mass in the region B_{UV}. Hence

$$F_{UV}(u, v) = \iint_{B_{UV}} f_{X,Y}(x, y)\, dx\, dy. \qquad \text{...(1)}$$

[We interpret $f_{X,Y}(x, y)$ as surface mass density in the x-y plane]

Joint Density Functions. The p.d.f. $f_{UV}(u, v)$ is given by

$$f_{U,V}(u, v) = \sum_i f_{X,Y}(x_i, y_i) \left| \frac{\partial(x_i, y_i)}{\partial(u, v)} \right| \qquad \text{...(2)}$$

where (x_i, y_i), $i = 1, 2, \ldots$ are the *real* solutions of $u = g(x, y)$ and $v = h(x, y)$.

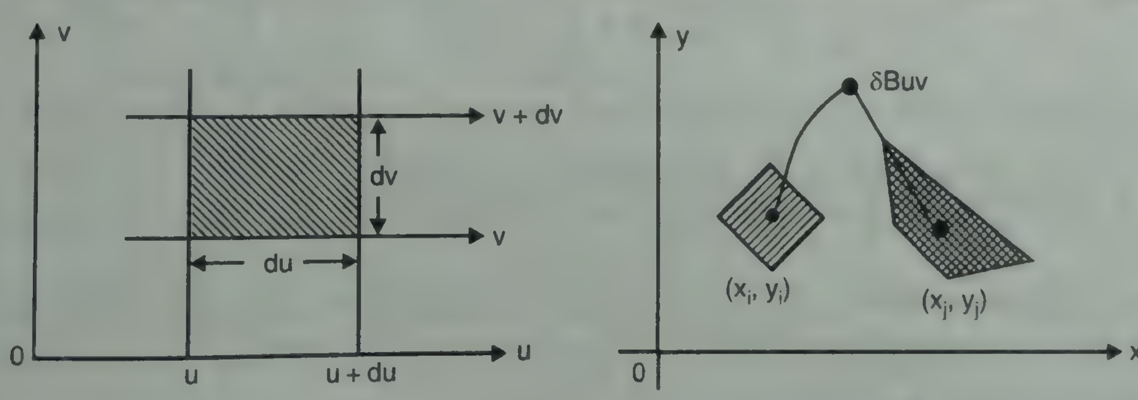
Proof. Let $\delta\beta_{UV}$ be the region of points in the x-y plane such that

$u < g(x, y) \le u + du$, $v < h(x, y) \le v + dv$.

Obviously, $\{u < U \le u + du, v < V \le v + dv\} = \{(X, Y) \in \delta\beta_{UV}$ ...(i)

But, $P(u < U \le u + du, v < V \le v + dv) = f_{U,V}(u, v)\, du dv$ ...(ii)

As we see from (i), the probability in (ii) equals the masses in the region $\delta\beta_{UV}$. Hence



$$f_{U,V}(u, v)\, du\, dv = \iint f_{X,Y}(x, y)\, dx\, dy \qquad \text{...(iii)}$$

Given v, u the region $\delta\beta_{UV}$ consists of differential parallelograms, one for each solution (x_i, y_i) of $u = g(x, y)$ and $v = h(x, y)$. From Calculus $dxdy = | \partial(x, y)/\partial(u, v)\, du\, dv$ hence the area of ith parallelogram is $dx_i dy_i$ and it contains the probability mass

$$f_{X,Y}(x_i, y_i)\, dx_i dy_i = f_{X,Y}(x_i, y_i) | \partial(x_i, y_i) / \partial(u, v) | du dv = f_{X,Y}(x_i, y_i) | J_i | du\, dv.$$

Summing the probability masses in all parallelograms, we obtain through (iii), the result

$$f_{UV}(u, v)\, du dv = \left\{ \sum_i \left| \frac{\partial(x_i, y_i)}{\partial(u, v)} \right| f_{X,Y}(x_i, y_i) \right\} du\, dv = \sum_i | J_i | f_{X,Y}(x_i, y_i)\, du dv.$$

From this result Eq. (2) follows instantly.

4-61. *Distribution of Sum, Difference, Product and Quotient via Jacobians*

In each of the following cases, we use an *auxiliary* r.v.W.

(A) Let $Z = aX + bY$. Put $z = ax + by$, $w = x$, then $|\partial(z, w)/\partial(x, y)| = b$.

From : $g_{Z.W}(z, w) = f(x, y).|J| = f\left(w, \frac{z - aw}{b}\right).\frac{1}{|b|}$, integrating **out** w we obtain.

$$g_Z(z) = \int_{-\infty}^{\infty} f\left(w, \frac{z - aw}{b}\right).\frac{dw}{|b|} \quad [a = b = 1 \Rightarrow Z = X + Y, -a = b = 1 \Rightarrow Z = Y - X]$$

$$g_S(z) = \int_{-\infty}^{\infty} f(w, z - w)\, dw, \qquad g_D(z) = \int_{-\infty}^{\infty} f(w, w - z)\, dw.$$

(*B*) Let $Z = XY$. Put $z = xy$, $w = x$, then $|\partial(z, w)/\partial(x, y)|\, x| = |w|$.

From : $g_{Z.W}(z, w) = f(x, y)\left|\frac{\partial(x, y)}{\partial(u, v)}\right| = f(w, z/w).\frac{1}{|w|}$; we integrate out w.

$$g_Z(z) = \int_{-\infty}^{\infty} f(w, z/w)\frac{dw}{|w|}.$$

(*C*) Let $Z = X/Y$. Put $z = x/y$, $w = y$, then $|\partial(x, y)/\partial(z, w)| = |w|$.

From : $g_{Z.W}(z, w) = f(x, y)\left|\frac{\partial(x, y)}{\partial(z, w)}\right| = f(wz, w).|w|$, we integrate **out** w

$$g_Z(z) = \int_{-\infty}^{\infty} f(wz, w)|w|\, dw.$$

Cor. Let X and Y be *independent*. Let $S = X + Y$, $D = Y - X$, $P = XY$, $Q = Y/X$. Then

$$\boxed{\begin{aligned}
g_S(z) &= \int_{-\infty}^{\infty} f(x, z - x)\, dx = \int_{-\infty}^{\infty} f(z - y, y)\, dy = \int_{-\infty}^{\infty} f_1(x)\, f_2(z - x)\, dx = \int_{-\infty}^{\infty} f_1(z - y)\, f_2(y)\, dy \\
g_D(z) &= \int_{-\infty}^{\infty} f(x, z + x)\, dx = \int_{-\infty}^{\infty} f(z + y, y)\, dy = \int_{-\infty}^{\infty} f_1(x)\, f_2(z + x)\, dx = \int_{-\infty}^{\infty} f_1(z + y)\, f_2(y)\, dy \\
g_P(z) &= \int_{-\infty}^{\infty} f\left(x, \frac{z}{x}\right)\frac{dx}{|x|} = \int_{-\infty}^{\infty} f\left(\frac{z}{y}, y\right)\frac{dy}{|y|} = \int_{-\infty}^{\infty} f_1(x)\, f_2\left(\frac{z}{x}\right)\frac{dx}{|x|} = \int_{-\infty}^{\infty} f_1\left(\frac{y}{z}\right) f_2(y)\frac{dy}{|y|} \\
g_Q(z) &= \int_{-\infty}^{\infty} f(x, zx)|x|\, dx = \int_{-\infty}^{\infty} f\left(\frac{y}{z}, y\right)\frac{|y|}{z^2}\, dy = \int_{-\infty}^{\infty} f_1(x)\, f_2(zx)|x|\, dx = \int_{-\infty}^{\infty} f_1\left(\frac{y}{z}\right) f_2(y)\frac{|y|}{z^2}\, dy
\end{aligned}}$$

4-62. *Worked-out Problems*

***Example* 1.** Let X and Y be i.i.d. $N(0, 1)$ variates. Suppose that

$$X = R\cos\theta,\ Y = R\sin\theta,\ 0 < R < \infty,\ 0 < \theta < 2\pi$$

Show that

(i) R ~ Rayleigh and θ ~ unif $(0, 2\pi)$ are independen..

(ii) R^2 ~ expo $(\frac{1}{2})$ and θ ~ unif $(0, 2\pi)$ are independent.

[Of course, (ii) follows from (i) by a simple map]

Solution. The joint probability differential of X, Y is

$$f(x, y)\, dx\, dy = f_1(x)\, dx.\quad f_2(y)\, dy = \frac{e^{-(x^2+y^2)/2}}{2\pi} dx\, dy \qquad \ldots(1)$$

(i) Let $x = r\cos\theta$, $y = r\sin\theta$, $|\partial(x, y)/\partial(r, \theta)| = r$ and (1) reduces to

$$g(r, \theta)\, d\theta\, dr = \left(\frac{1}{2\pi} d\theta\right)(re^{-r^2/2} dr), \qquad 0 < \theta < 2\pi,\ 0 < x < \infty.$$

This shows that R and θ are independent distributed with stated densities.

(ii) For convenience, we write $W = R^2$. Then set $x = \sqrt{w}\cos\theta$, $y = \sqrt{w}\sin\theta$

$$\frac{\partial(x, y)}{\partial(w, \theta)} = \begin{vmatrix} \cos\theta / 2\sqrt{w} & \sin\theta / 2\sqrt{w} \\ -\sqrt{w}\sin\theta & \sqrt{w}\cos\theta \end{vmatrix} = \frac{1}{2}$$

The joint probability differential (1) now reduces to

$$h(w, \theta)\, d\theta\, dw = \left(\frac{1}{2} e^{-w/2} dw\right)\left(\frac{d\theta}{2\pi}\right), \qquad 0 < w < \infty,\ 0 < \theta < 2\pi.$$

This shows that $W = R^2$ and θ are independently distributed with stated densities.

***Example* 2.** Suppose $X = \sqrt{w}\cos\theta$, $Y = \sqrt{w}\sin\theta$. Let $W = R^2 \sim$ expo $(\frac{1}{2})$ and $\theta \sim$ unif $(0, 2\pi)$ be independent.

Show that X and Y are i.i.d. $N(0, 1)$ variates.

Solution. The joint probability differential of W and θ is

$$f(w, \theta)\, dw d\theta = \left(\tfrac{1}{2} e^{-w/2} dw\right)(d\theta / 2\pi) \qquad \ldots(2)$$

Now $x^2 + y^2 = w$, $\quad |\partial(x, y)/\partial(w, \theta)| = \frac{1}{2}$, hence $dw\, d\theta = 2dx\, dy$.

Substituting into (2) yields

$$g(x, y)\, dx\, dy = \frac{1}{2\pi} e^{-(x^2+y^2)/2} dx\, dy$$

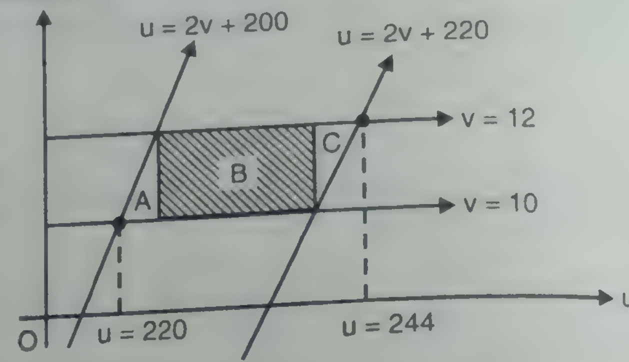
$$= \left(\frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx\right)\left(\frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy\right) \quad -\infty < x, y < \infty.$$

Thus X and Y are i.i.d. $N(0, 1)$ variates. [Converse to Example 1(ii)]

***Example* 3.** Independent variates X and Y, representing the length and breadth of a rectangle, have the densities

$$f_1(x) = 1/10,\ 100 < x < 110 : f_2(y) = \tfrac{1}{2},\ 10 \le y \le 12; \text{ vanish elsewhere.}$$

Find the p.d.f. of the perimeter of the rectangle.



Solution. Let $U = 2(X + Y)$, (perimeter); $V = Y$, so that $X = \frac{1}{2}U - V, Y = V$. Then $\partial(x, y)/\partial(u, v) = \frac{1}{2}$; $f(x, y) = 1/20$ transforms to $g(u, v) = 1/40$.

$S = \{(x, y) : 100 < x < 110; 10 \le y \le 12\}$ transforms to $T = \{(u, v) : 200 < u - 2v < 220 ; 10 \le v \le 12\}$

The parallelogram T in u-v plane can be split into three parts A, B, C as shown. Hence

$$g_1(u) = \int_{10}^{a} \frac{dv}{40} = \frac{a-10}{40} = \frac{u-220}{80}, \qquad 220 \le u \le 224$$

$$g_1(u) = \int_{10}^{12} \frac{dv}{40} = \frac{1}{2}, \qquad 224 \le u \le 240$$

$$g_1(u) = \int_{b}^{12} \frac{dv}{40} = \frac{12-b}{80} = \frac{224-u}{80}, \qquad 240 \le u \le 244.$$

Problems with Solutions Provided at the End of the Text

1*. The joint p.d.f. of variates X and Y is

$$f(x, y) = 4xye^{-(x^2+y^2)}, x \ge 0, y \ge 0; \quad f(x, y) = 0, \text{ elsewhere}$$

(a) Are X and Y independent ? Find the conditional density of X given $Y = y$.

(b) Prove that the density function of distance $u = (X^2 + Y^2)^{1/2}$ is

$$g(u) = 2u^3 e^{-u^2}, 0 \le u < \infty; \quad g(u) = 0, \text{ elsewhere.}$$

2*. Independent variates X and Y have the p.d.f.'s

$$f(x) = (\pi\sqrt{1-x^2})^{-1}, |x| < 1; \quad g(y) = ye^{-y^2/2}, \ y > 0.$$

Find the joint p.d.f. of $Z = XY$ and $W = X$. Hence or otherwise find the density of Z.

3*. Let $(X, Y) \sim$ Dirch (a, b, c). If $X = Z$ and $Y = T(1 - Z)$, find the distributions of T and Z. Let U, V, W be indep. $\gamma(\ell)$, $\gamma(m)$, $\gamma(n)$ variates.
If $P = U/(U + V + W)$, $Q = V/(U + V + W)$, then $(P, Q) \sim$ Dirch (l, m, n).

4*. Let $f(x, y) = k(x + y)$; $0 < x < 1$, $0 < y < 1$, $0 < x + y < 1$.
Find $f_X(x)$ and obtain the joint and marginal distributions of $X + Y$ and $X - Y$.

Exercise 4(c)

1. Find the p.d.f. of $X + Y$, when the joint p.d.f. of X and Y is

$$f(x, y) = \tfrac{1}{2}xe^{-y}, \ 0 < x < 2, y > 0, \ f(x, y) = 0; \text{ elsewhere.}$$

[**Ans.** $f(u) = \frac{1}{2}(u - 1 + e^{-u})\, I(0 < u \le 2) + \frac{1}{2}(1 + e^2)e^{-u} I(2 < u < \infty)$].

2. Find the p.d.f. of $Z = X/Y$, if X and Y are two independent random variables such that

$$f(x) = \lambda e^{-\lambda x}, x \ge 0; \ g(y) = \mu e^{-\mu y}, y \ge 0.$$

[**Ans.** $f(x) = (\mu/\lambda)\,[z + (\mu/\lambda^2]^{-2}, z > 0$]

3. If $f(x, y) = (y/x^2)\, I(|x| \geq 1)\,.\, I(0 \leq y \leq 1)$, show that the distribution of $Q = X/Y$ is
$f_Q(t) = (t^{-2} + t^{-3})\, I(t \leq -1) + (t^{-2} - t^{-3})\, I(t \geq 1)$.

4. Variates X and Y are jointly distributed as follows :

$$f(x, y) = 2(a^2 - x^2 - y^2)/\pi a^4, 0 \leq x^2 + y^2 \leq a^2.$$

(a) Show that the marginal distribution of X is

$$f_X(x) = 8(a^2 - x^2)^{3/2}/3\pi a^4, -a \leq x \leq a.$$

(b) Show that the joint prob. distribution of $r = \sqrt{x^2 + y^2}$ and $\theta = \tan^{-1}(y/x)$, is

$$p(r, \theta) = 2r(a^2 - r^2)/\pi a^4, 0 \leq r \leq a, 0 < \theta < 2\pi.$$

(c) Show that the p.d.f. of R is $h(r) = 4r(a^2 - r^2)/a^4, 0 \leq r \leq a$.

(d) Show that the joint distribution of X and R, where $R = \sqrt{X^2 + Y^2}$ is

$$g(x, r) = 4r(a^2 - r^2)/\pi\sqrt{(r^2 - x^2)}a^4,\ 0 < r < a,\ -r < x < r.$$

5. (i) Let X and Y be a random sample of size 2 from a distribution with p.d.f.

$f(x) = \lambda e^{-\lambda x}, 0 < x < \infty; f(x) = 0$, elsewhere.

Show that $X + Y$ and $X/(X + Y)$ are independent.

(ii) Let X, Y, Z denote a random sample of size 3 drawn from the above distribution. Show that $U = X + Y + Z$, $V = X/(X + Y)$, $W = (X + Y)/(X + Y + Z)$ are mutually independent.

6. If X and Y have the joint p.d.f. $f(x, y)$, find the joint p.d.f. of $Z = \sqrt{X^2 + Y^2}$ and $W = X/Y$.

7. The p.d.f. of (X, Y) is $f(x, y) = abe^{-ax - by}, x, y > 0; f(x, y) = 0$, elsewhere.

Find the density function of (U, V) where $U = 1 + 2(X/Y)$, $V = 3X + Y$.

$$\textbf{[Ans. } g(u, v) = \frac{2abv}{(3u - 1)^2} \exp\left[-\left(\frac{au + 2b - a}{3u - 1}\right)v\right] I(v > 0)\ I(u > 1)$$

$$g(u) = [2ab/(au + 2b - a)^2]\ I(u > 1)]$$

8. Let X_1 and X_2 be two independent determinations of a variate X whose density is

$$f(x) = a/x^2, x > a; f(x) = 0, \text{elsewhere}$$

Show that the p.d.f. of $Y = X_1/X_2$ is $g(y) = (\frac{1}{2})\, I(0 \leq y \leq 1) + (2y^2)^{-1}\, I(y \geq 1)$.

9. If $f(x, y) = ye^{-xy}\, I(0 < x < \infty)\, I(0 < y < 1)$, find the distribution of (i) $U = \min(X, Y)$ and (ii) $V = \max(X, Y)$.

$$\textbf{[Ans. } f(u) = [e^{-u^2}(2 + u^{-2}) - e^{-u}(u^{-2} + u^{-1})]\, I(0 < u < 1), F(t) = 0, I(t \leq 0) + [t + t^{-1}(e^{-t^2} - 1)]$$

$$I(0 < t < 1) + [1 - t^{-1}(1 - e^{-t})]\, I(t \geq 1)]$$

10. If $f(x, y) = 2e^{-(x+y)}\, I(0 \leq y < x < \infty)$, find the distribution of $S = X + Y$ and $Q = X/Y$.

$$\textbf{[Ans. } f_S(z) = ze^{-z}\, I(z > 0)\,.\, f_Q(t) = 2(1 + t)^{-2}\, I(t > 1)]$$

11. Variates X and Y have the joint p.d.f. $f(x, y)$, given by

$f(x, y) = g(x + y), x > 0, y > 0;\ f(x, y) = 0$, otherwise.

Obtain the c.d.f. $H(z)$ of $Z = X + Y$ and deduce that its p.d.f. is

$$h(z) = zg(z), z > 0;\ h(z) = 0,\ z \leq 0.$$

4-70. Miscellaneous Worked-out Problems

***Example* 1.** (i) Show that even if two random variables are not independent, their functions still can be independent.

(ii) Give an example of two r.v.'s X and Y that are not independent but such that X^2 and Y^2 are independent.

Solution. Let $\Omega = \{a, b, c, d\}$, where all the outcomes are euqiprobable. Define variates X and Y on Ω by the relations

$$X(a) = 1,\ X(b) = 0,\ X(c) = -1,\ X(d) = 0;\ Y(a) = 1,\ Y(b) = -1,\ Y(c) = 0 = Y(d).$$

Then $P(X = 1) = \frac{1}{4} = P(Y = 1)$ each $P(\{a\})$; and $P(X = 1, Y = 1) = P(\{a\}) = \frac{1}{4}$.

Obviously : $P(X = 1, Y = 1) \neq P(X = 1)\ P(Y = 1) \Rightarrow X$ and Y not indep.

Now consider the distributions of X^2 and Y^2. Let $U = X^2$, $V = Y^2$, then

$$P(U = 0) = \tfrac{1}{2} = P(U = 1);\ P(V = 0) = \tfrac{1}{2} = P(V = 1).$$

Now $\quad P(U = 1, V = 1) = P(a) = \frac{1}{4} = P(U = 1)\ P(V = 1),$

$$P(U = 0, V = 0) = P(d) = \tfrac{1}{4} = P(U = 0)\ P(V = 0),$$

$$P(U = 0, V = 1) = P(b) = \tfrac{1}{4} = P(U = 0)\ P(V = 1),$$

$$P(U = 1, V = 0) = P(c) = \tfrac{1}{4} = P(U = 1)\ P(V = 0).$$

This shows that U and V are independent, i.e. X^2 and Y^2 are independent.

(ii) Let $f(x, y) = (1 + xy)/4,\ -1 < x, y < 1;\ f(x, y) = 0$, elsewhere

Then $\quad f_1(x) = \frac{1}{4}\int_{-1}^{1}(1 + xy)\,dy = \frac{1}{2}, |x| < 1.$ Similarly, $f_2(y) = \frac{1}{2},\ |y| < 1$

Since $f(x, y) \neq f_1(x)\,f_2(y)$, it follows that X and Y are not independent. However,

$$P\{X^2 \le a, Y^2 \le b\} = \int_{-\sqrt{a}}^{\sqrt{a}}\int_{-\sqrt{b}}^{\sqrt{b}} f(x, y)\,dx\,dy = a^{1/2}.b^{1/2} = P(X^2 \le a)\,.\,P(Y^2 \le b^2)$$

It follows that X^2 and Y^2 are independently distributed.

***Example* 2.** Let $F(t)$ be a cumulative distribution function (c.d.f.).

(i) Is $G(x, y) = F(x) + F(y)$, a joint c.d.f. ?

(ii) Is $G(x, y) = \max\{F(x), F(y)\}$, a joint c.d.f. ?

(iii) Is $G(x, y) = F(x).\ F(y)$, a joint c.d.f. ?

(iv) Is $G(x, y) = \min\{F(x), F(y)\}$, a joint c.d.f. ?

***Solution. Part* (i) :** Here $\quad G(x, y) = F(x) + F(y)$

$$\therefore \quad G(-\infty, y) = F(-\infty) + F(y) \Rightarrow G(-\infty, \infty) = F(-\infty) + F(\infty) = 0 + 1 = 1$$

Since $G(-\infty, \infty) = 1 \neq 0$, [Art. 4-11(4)], $G(x, y)$ cannot be a joint c.d.f.

***Part* (ii) :** Here $G(x, y) = \max\{F(x), F(y)\}$

$$G(-\infty, \infty) = \max\{F(-\infty)\}, F(\infty) = \max\{0, 1\} = 1$$

Since $G(-\infty, \infty) = 1 \neq 0$, we conclude that $G(x, y)$ is not a joint c.d.f.

Part (iii) : Here $G(x, y) = F(x) . F(y)$.

Step 1. Since $F(t)$ is a c.d.f. so $0 \le F(t) \le 1$ for all t and hence $0 \le F(x) F(y) \le 1$, $\forall\ x, y$

i.e. $$0 \le G(x, y) \le 1. \quad ...(i)$$

Step 2. Since $F(t) \ge 0$ is monotinically increasing (for all t), hence $F(x) F(y) \ge 0$ is also monotonically increasing. Thus $G(x, y)$ is monotonically increasing in both arguments.

Step 3. $G(-\infty, y) = F(-\infty) F(y) = 0$; $G(x, -\infty) = F(x) F(-\infty) = 0$; $G(\infty, \infty) = F(\infty) F(\infty) = 1$. ...(iii)

Step 4. We know that $F(t)$ is right continuous, hence

$$\lim_{x \to a+} G(x, y) = \lim_{x \to a+} [F(x) . F(y)] = [\lim_{x \to a+} F(x)] F(y) = F(a) F(y) = G(a, y).$$

Hence $G(x, y)$ is right continuous in its first argument. Similarly, $G(x, y)$ is right-continuous in its 2nd argument. Thus, $G(x, y)$ is right continuous in both arguments.

Step 5. Now consider the *Rectangle Rule :*

$$\begin{aligned} p &= G(a, c) - G(a, d) - G(b, c) + G(b, d), \qquad [a < b, c < d] \\ &= F(a) F(c) - F(a) F(d) - F(b) F(c) + F(b) F(d) \\ &= [F(b) - F(a)] [F(d) - F(c)] \end{aligned}$$

Since $F(x)$ is increasing, so $a < b$, $c < d$ gives $F(b) > F(a)$, $F(d) > F(c)$, hence $p \ge 0$.

Also $F(b) - F(a) \le F(b) \le 1$; $F(d) - F(c) \le F(d) \le 1$, so $p \le 1$.

Combining these results ; $0 \le p \le 1$.

These 5 steps establish that $G(x, y)$ is indeed a c.d.f.

Part (iv) : Here $G(x, y) = \min \{F(x), F(y)\}$.

Step 1. Since $0 \le F(x), F(y), \le 1 \Rightarrow 0 \le \min \{F(x), F(y)\} \le 1 \Rightarrow 0 \le G(x, y) \le 1$.

Step 2. Let $x_1 > x_2$, then $F(x_1) > F(x_2)$ because c.d.f. 'F' is always increasing.

$$G(x_1, y) = \min \{F(x_1), F(y)\}; G(x_2, y) = \min \{F(x_2), F(y)\}. \quad (A)$$

The following three cases arise :

(a) $F(x_1) \ge F(y) \ge F(x_2)$, (b) $F(x_1) \ge F(x_2) \ge F(y)$, (c) $F(y) \ge F(x_1) \ge F(x_2)$.

For the case (a) : $G(x_1, y) = F(y)$, $G(x_2, y) = F(x_2) \Rightarrow G(x_1, y) \ge G(x_2, y)$.

For the case (b) : $G(x_1, y) = F(y)$, $G(x_2, y) = F(y) \Rightarrow G(x_2, y) = G(x_1, y)$.

For the case (c) : $G(x_1, y) = F(x_1)$, $G(x_2, y) = F(x_2) \Rightarrow G(x_1, y) \ge G(x_2, y)$.

Hence $G(x, y)$ is monotonic in x-coordinate ; similarly it is monotonic for y-coordinate. That is $G(x, y)$ is monotonic in both arguments x and y.

Step 3. $G(-\infty, y) = \min \{F(-\infty), F(y)\} = 0$, $[\because F(-\infty) = 0]$. Similarly, $G(x, -\infty) = 0$.

Also $G(\infty, \infty) = \min \{F(\infty), F(\infty)\} = 1$, since $F(\infty) = 1$.

Step 4. Since $F(x)$ is a c.d.f., it is right continuous. So using definition of minimum;

$$G(x, y) = \min \{F(x), F(y)\} = \tfrac{1}{2}[F(x) + F(y)] - [\lim_{x \to \infty} F(x) - F(y)|],$$

$$\therefore \lim_{x \to a+} G(x, y) = \tfrac{1}{2} \lim_{x \to a+} [F(x) + F(y) - [F(x) - F(y)|] = \tfrac{1}{2}[\lim_{x \to a+} F(x) + F(y) - |F(x) - F(y)|],$$

$$= \tfrac{1}{2}[F(a) + F(y) - |F(a) - F(y)|] = \min [F(a), F(y)] = G(a, y).$$

Thus $G(x, y)$ is right continuous in x-coordinate. Since same argument holds for y-coordinate, we infer that $G(x, y)$ is right continuous.

***Step* 5.** Now consider the following Rectangle Rule for distribution functions.

$$p = G(a, c) - G(a, d) - G(b, c) + G(b, d), \text{ for } a < b,\ c < d. \qquad \ldots(1)$$

The following six cases arise :

(i) $a < b < c < d$ (ii) $c < a < b < d$ (iii) $c < a < d < b$

(iv) $c < d < a < b$ (v) $a < c < d < b$ (vi) $a < c < b < d$.

It suffices to consider any one case, since all other cases behave similarly. Suppose we consider (v), then $a < c < d < b$. Since $F(t)$ is monotonic increasing, we have

$$F(a) \le F(c) \le F(d) \le F(b)$$

Since $G(x, y) = \min\{F(x), F(y)\}$, use of (2) into (1) provides

$$p = \min\{F(a), F(c)\} - \min\{F(a), F(d)\} - \min\{F(b), F(c)\} + \min\{F(b), F(d)\}$$

$$= F(a) - F(a) - F(c) + F(d) = F(d) - F(c) \ge 0, \qquad \text{[by (2)]}$$

Thus $\quad p = F(d) - F(c) \le F(d) \le 1$, and hence $0 \le p \le 1$.

All other cases lead to $0 \le p \le 1$. We conclude that $G(x, y)$ is a joint c.d.f.

Problems with Solutions Provided at the End of the Text

1*. The following table gives the joint p.d.f of X and Y :

$x \downarrow$ \ $y \rightarrow$	y_1	y_2	y_3	y_4	$P(X = x)$
x_1	$2a$	$2a$	$2a$	$10a$	$16a$
x_2	$2a$	$2a$	a	$11a$	$16a$
x_3	$8a$	$8a$	$9a$	$39a$	$64a$
$P(Y = y)$	$12a$	$12a$	$12a$	$60a$	1

Are X and Y independent variables ?

2*. Let X and Y be independent variables with distribution :

$$P(X = \pm 1) = \tfrac{1}{2} = P(Y = \pm 1). \text{ Are } X \text{ and } Z = XY \text{ independent ?}$$

3*. Let $f(x, y) = k(y - x)^{\beta}$, $0 \le x \le y \le 1$; $f(x, y) = 0$, elsewhere,

(i) For what values of β can k be chosen to make f a p.d.f. ?

(ii) How should k be chosen to make f a p.d.f. ?

(iii) Find the individual densities of f.

4*. If $f(x, y) = 1$, $0 < x < 1$, $0 < y < 1$, find the density of Z, where $Z = X + Y$, $x + y \le 1$; $Z = X + Y - 1$, $x + y \ge 1$.

5*. The joint distribution of X and Y is $f(x, y) = 2e^{-(x+y)}$, $0 < x < y$, $0 < y < \infty$. Find the joint and marginal distributions of X and $X + Y$.

6*. Given $f(x, y \mid z) = z + (1 - z)(x + y), 0 < x, y < 1, 0 \le z \le 2$ and $f_Z(z) = \frac{1}{2}, 0 \le z \le 2$.

(a) Are X, Y independent ?

(b) Are X, Z independent ?

(c) Find the joint distribution of X and $X + Y$.

(d) Find the distribution of $\max\{(X, Y) \mid Z = z\}$.

(e) Find the distribution of $\{(X + Y) \mid Z = z\}$.

Miscellaneous Exercises

1. Show why the function f defined by

$f(x, y) = kx(3x - y), 0 < x < 2, -x < y < 4x$; $f(x, y) = 0$, elsewhere

cannot represent a p.d.f. for any choice of k.

2. The joint p.d.f. of (X, Y) is

(a) $f(x, y) = \frac{3}{2}(2 - 2x - y), x = 0, y = 0, y \le 2 - 2x$. Find $f_1(x)$ and $f_2(y)$.

(b) $f(x, y) = x^2 + \frac{1}{3}xy, 0 \le x \le 1, 0 \le y \le 2$. Find $P\{X > \frac{1}{2} \mid Y > \frac{1}{2}\}, P\{Y < \frac{1}{2} \mid X < \frac{1}{2}\}, P\{X + Y < 1\}$.

(c) $F(x, y) = 8xy, 0 < x < y < 1$. Find the joint density of X/Y and Y.

[**Ans.** (a) $f_1(x) = 3(1 - x)^2 I(0 < x < 1), f_2(y) = [3(2 - y)^2/8] I(0 < y < 2)$.

(b) $p = 43/52$, (c) $f(u, v) = f_1(u) . f_2(v) = (2u)(4v^3), 0 < u, v < 1$]

3. The joint p.d.f. of X and Y is given by : $f(x, y) = \frac{1}{2}xy, 0 < y < x, \; 0 < x < 2$.

Find the marginal distributions of X and Y. Are X and Y independent ?

[**Ans.** $f_1(x) = (x^3/4) I(0 < x < 2), f_2(y) = [y - (y^3/3)] I(0 < y < 2)$]

4. A hand of three cards is dealt with from an ordinary deck of 52 cards. If X and Y are the number of aces and the number of spades in hand, find the joint p.d.f. of (X, Y).

5. If $f(x, y) = (2y/x^2) I(x \ge 1) I(0 \le y \le 1)$, find the distribution of $S = X + Y$ and $T = XY$.

[**Ans.** $f_S(t) = 2(t - 1 - \ln t) \; I(1 \le t < 2) + 2[(t - 1)^{-1} - \ln t \,(t - 1)^{-1}] \, I(t \ge 2)$

$F_T(f) = (2t/3) I(0 < t < 1) + (2/3t^2) I(t \ge 1)$]

6. Four coins are tossed. Let X be the number of heads and Y be the number of heads minus the number of tails. Find the p.d.f. of X, the p.d.f. of Y and show that $P(-2 \le Y < 4) = 7/8$.

7. Let (X, Y) have the density $f(x, y) = \frac{1}{2}$, for (x, y) inside the square with corners $(a, a), (-a, a), (a, -a), (-a, -a)$ and $f(x, y) = 0$, elsewhere. Find a and obtain the marginal densities for X and Y. How the values of a and marginal densities change if the corners of the square are $(a, 0), (-a, 0), (0, a), (0, -a)$.

[**Ans.** $f_1(t) = f_2(t) = a = 1/\sqrt{2} \; |t| < a; \; g_1(t) = g_2(t) = 1 - |t|, |t| < 1.$]

8. Let X_1, X_2 be a random sample from a distribution with p.d.f. $f(x) = 2x, 0 < x < 1$; zero, elsewhere. Show $P[(X_1/X_2) \le \frac{1}{2}] = \frac{1}{8}$, and $P\{X_1 < X_2 \mid X_1 < 2X_2] = \frac{1}{4}$

9. Obtain the marginal d.f.s. of X and Y, when joint d.f. of X and Y is given as follows :

$$F(x, y) = \begin{cases} 0, & x < 0 \quad \text{or} \quad y < 0 \\ \frac{1}{2}(1 - \cos y + \sin y), & x \geq \pi/2; \quad 0 \leq y < \pi/2 \\ \frac{1}{2}(1 - \cos x + \sin x), & 0 \leq x < \pi/2; \quad y \geq \pi/2 \\ \frac{1}{2}[\sin x + \sin y - \sin(x+y)], & 0 \leq x < \pi/2; \quad 0 \leq y < \pi/2 \\ 1, & x \geq \pi/2; \quad y \leq \pi/2 \end{cases}$$

$\left[\textbf{Ans. } F_1(t) = F_2(t) = 0 I(t < 0) + \frac{1}{2}(1 - \cos t + \sin t)\, I(0 \leq t < \pi/2) + 1\, I(t \geq \pi/2)\right]$

10. Let $f(x \mid y) = kx/y^2,\ 0 < x < y,\ 0 < y < 1;\ f(x \mid y) = 0$, elsewhere and $f_Y(y) = cy^4,\ 0 < y < 1;\ f_Y(y) = 0$, elsewhere.
   (a) Show that $k = 2$ and $c = 5$ and find the joint p.d.f. of X and Y.
   (b) Prove that $P\left[\frac{1}{4} < X < \frac{1}{2} \mid Y = \frac{5}{8}\right] = \frac{12}{25}$, and $P\left[\frac{1}{4} < X < \frac{1}{2}\right] = \frac{449}{1536}$.

11. Find the marginal and conditional p.d.f.'s if the joint density is
   (a) $f(x, y) = 2(2 - x - y),\ 0 \leq x \leq y \leq 1;\ f(x, y) = 0$, elsewhere.
   (b) $f(x, y) = \dfrac{9(1 + x + y)}{2(1 + x)^4 (1 + y)^4},\ 0 \leq x, y < \infty.$

[**Ans.** $f_1(x) = 3(1 - x^2)\, I(0 \leq x \leq 1), f_2(y) = (4y - 3y^2)\, I(0 < y < 1)$,
(b) $f_1(t) = f_2(t) = [3(3 + 2t)/4(1 + t)^4]\, I(t > 0); f(x \mid y) = [6(1 + x + y)/(3 + 2y)(1 + x)^4]$]

12. Obtain the conditional distributions when variates X and Y have the joint distributions :

(i)

$y \downarrow$ \ $x \rightarrow$	1	2	3
1	$3a$	$6a$	0
2	0	$4a$	$3b$
3	$2a$	$9a$	$2b$

(ii)

$y \downarrow$ \ $x \rightarrow$	1	2	3
1	1/12	1/6	0
2	0	1/9	1/5
3	1/18	1/4	2/15

where $12a = 5b$.

13. Variates X and Y have the joint distribution given in the following table. Find a.

$y \downarrow$ \ $x \rightarrow$	−3	0	2	4
−4	a	0	$2a$	a
−3	0	a	0	$3a$
5	$2a$	0	a	a

Show that $P(X \geq Y) = \frac{5}{12}$, $P(X^2 > Y^2) = \frac{2}{3}$, $P(X + Y < 0) = \frac{1}{4}$ and $P(XY > 0) = \frac{1}{2}$.

14. Show that the conditions for the function $f(x, y) = k \exp(ax^2 + 2hxy + by^2),\ -\infty < x, y < \infty$ to be a bivariate p.d.f. are : $a \leq 0, b \leq 0, ab - h^2 \geq 0$.
Show that, when these conditions are satisfied, $\pi k = (ab - h^2)^{1/2}$.

[**Hint.** Let $x = r\cos\theta$, $y = r\sin\theta$; then $\partial(x, y)/\partial(r, \theta) = r$. Thus
$f(r, \theta) = kr\exp(Br^2)$, $0 \le r < \infty, 0 < \theta < 2\pi$.
$B = a\cos^2\theta + 2h\sin\theta\cos\theta + b\sin^2\theta = (1/a)\,[(a\cos\theta + h\sin\theta)^2 + (ab - h^2)\sin^2\theta]$
$= (1/b)\,[b\sin\theta + h\cos\theta)^2 + (ab - h^2)\cos^2\theta]$.

The Improper integral $\int_0^\infty re^{Br^2}\,dr$ is convergent iff $B < 0$.

This holds iff $a < 0$, $b < 0$ and $ab - h^2 > 0$. Put $B = -\lambda$, $(\lambda > 0)$.
Use normalization to obtain the value of k; thus

$$I = k\int_0^{2\pi} d\theta \int_0^\infty re^{-\lambda r^2}\,dr = \frac{k}{2}\int_0^{2\pi} \frac{d\theta}{a\cos^2\theta + 2h\sin\theta\cos\theta + b\sin^2\theta}.$$

Put $\tan\theta = z$ and evaluate. The cases $a = 0$, $b = 0$, $ab = h^2$ may be discussed separately]

15. Two discrete variates X and Y have the p.d.f.
(i) $p(0, 0) = 2/9$, $p(0, 1) = 1/9$, $p(1, 0) = 1/9$, $p(1, 1) = 5/9$. [$p(x, y) = P(X = x, Y = y)$]
Show that X and Y are not independent.

(ii) $p(0, 0) = p(0, 1) = p(1, 2) = \frac{1}{3}$. Find F_X, $F_Y F(x, y)$. Show that point $(-\frac{1}{2}, 0)$ is a continuity point and point $(0, 1)$ is a discontinuity point of $F(x, y)$.

16. Suppose X and Y are two variates where Y is *degenerate*. Show that X and Y are independent.

17. If X and Y are independnet, absolutely continuous variates, show that their joint distribution is also absolutely continuous.

18. Let X and Y be independent variatets with $P(X \ge 0) = 1 = P(Y \ge 0)$.
Are (i) $P(XY \le 6) = P(X \le 2)\,P(Y \le 3)$, (ii) $P(XY < 6) \ge P(X \le 2)\,P(Y \le 3)$ true ?

19. Consider two events A and B such that : $P(A) = \frac{1}{4}$, $P(B \mid A) = \frac{1}{2}$, $P(A \mid B) = \frac{1}{4}$. Define the variates X and Y as under :
$X = 1$, if A has occurred ; $X = 0$, if A has not occurred ;
$Y = 1$, if B has occurred ; $Y = 0$, if B has not occurred ;
Now comment on the following statements :
(a) X and Y are independent (b) $P(X^2 + Y^2 = 1) = \frac{1}{4}$, $P\{XY = X^2Y^2\} = 1$.

20. The joint p.d.f. of X and Y is : $f(x, y) = xe^{-x(1+y)}$, $x > 0$, $y > 0$; $f(x, y) = 0$, elsewhere.
Find (i) the p.d.f. of $Z = XY$, (ii) c.d.f. of $W = \max(X, Y)$.
[**Ans.** $f(z) = e^{-z}$, $z > 0$; $F(w) = (1 - e^{-w}) + (1 + w)^{-1}\,[e^{-w-w^2} - 1]$, $w > 0$]

21. If $f(x, y) = 4xy$, $0 < x, y < 1$, find the joint density of X^2 and Y^2.

22. (a) Variates X, Y, Z have the joint p.d.f.
$f(x, y, z) = 12x^2yz$, $0 < x, y, z < 1$; $f(x, y, z) = 0$, elsewhere.
Find marginal densities of X, Y, Z and the joint p.d.f. of X and Y.
What is the conditional p.d.f. of X and Y given $Z = z_0$, $0 < z_0 < 1$.
Also, show that $P(X \le Y) = \frac{2}{5}$, $P(X < Y < Z) = \frac{4}{35}$.

(b) If $f(x, y, z) = 8xyz$, $0 < x, y, z < 1$, show that $P(X < Y < Z) = \frac{1}{6}$.

23. Let X, Y, Z be i.i.d. variates from a distribution having p.d.f. $f(x) = 5x^4$, $0 < x < 1$, $f(x) = 0$, elsewhere. If $W = \max(X, Y, Z)$, show that $F(w) = w^{15}\, I(0 < w < 1)$ and find p.d.f. of W.

24. If $f(x, y) = e^{-(x+y)}$, $0 < x, y < \infty$; show that the joint density of X and $X + Y$ is $g(u, v) = e^{-u}$, $0 < v < u < \infty$.

25. If $f(x, y) = 2(x + y - 3xy^2)\, I(0 < x < 1)\, I(0 < y < 1)$, find the distribution of $M = \max(X, Y)$ and $N = \min(X, Y)$.

[**Ans.** $f_M(t) = \frac{1}{2}(9t^2 - 5t^4)\, I(0 < t < 1)$, $f_N(t) = (2 + 2t - 9t^2 + 5t^4)\, I(0 < t < 1)$]

26. If $f(x, y, z) = e^{-(x+y+z)}$, $0 < x, y, z < \infty$; show that the density of $W = \frac{1}{3}(X + Y + Z)$ is $g(w) = (27/2)w^2 e^{-3w}$, $w > 0$.

27. Prove or disprove :

(a) If $Y = X + 1$, then $F_Y(t) = F_X(t + 1)$ for all t.

(b) If $F_X(t) > F_Y(t)$ for all t, then $P(X < Y) > 0$.

28. The function $\phi(x)$ is monotone increasing and $Y = \phi(X)$. Prove that $F(x, y) = F_1(x)\, I(y > \phi(x)) + F_2(y)\, I(y < \phi(x))$. What happens if $\phi(x)$ is monotone decreasing ?

To live for some future goal is shallow. It's the sides of the mountain that sustain life, not the top. [Robert Pirsig]

Mathematical Expectation 5

5-10. Mathematical Expectation

1. Let X be a discreate random variable defined on a probability space, and let $\{x_i\}$ be the countable set of its possible values, such that $P(X = x_i) = p_i$. The mathematical expectation of X, (or the *mean* of X), written $\mathrm{E}(X)$, is a real number defined by the kernel-sum

$$\mathrm{E}(X) = \sum p_i x_i, \text{ for all } i \qquad \text{(kernel definition)} \qquad ...(1)$$

provided that the series $\sum p_i \,|\, x_i \,|$ is convergent.

2. Let X be a continuous r.v. with p.d.f. $f(x)$. If the integral

$$\int_{-\infty}^{\infty} f(x) |\, x \,|\, dx < \infty$$

(i.e. convergent), the expectation of X, written $\mathrm{E}(X)$, is the real number defined by

$$\mathrm{E}(X) = \int_{-\infty}^{\infty} f(x)\, x\, dx. \qquad ...(2)$$

Comments. If the series (1) or integral (2) is conditionally convergent, $\mathrm{E}(X)$ does not exist. The absolute convergence of the defining sum (1) or integral (2) for the existence of $\mathrm{E}(X)$ is pre-requisite. Thus, $\mathrm{E}(X)$ exists iff $\mathrm{E}(|\, X \,|)$ exists.

Note : If the values x_j (and hence the corresponding p_j as well) are listed in a different order, the infinite sum $\sum p_j x_j$ may change its value although the r.v. X remains the same. Hence it is but natural that we insist on the *absolute* convergence so that rearrnagement of series $\sum p_j x_j$ does not alter its value (sum), yielding unique value of $\mathrm{E}(X)$.

Warning Problem. A r.v. X takes the value $x_k = (-1)^{k-1} 2^k / k, 1 \le k < \infty$, with probability $p_k = 2^{-k}$. Find $\mathrm{E}(X)$.

Solution. We emphasize that $\mathrm{E}(X) < \infty$ iff $\mathrm{E}\,|\, X \,|$ exists. Here

$$\mathrm{E}(|\, X \,|) = \sum_{k=1}^{\infty} \frac{2^k}{k} . 2^{-k} = \sum_{k=1}^{\infty} \frac{1}{k} = 1 + \frac{1}{2} + \frac{1}{3} + \ldots$$

This is a well-known harmonic series which is divergent. Hence $\mathrm{E}(X)$ does not exist Let us note the blind-folded attack :

$$\mathrm{E}(X) = \sum (-1)^{k-1} (2^k / k)(2^{-k}) = \sum [(-1)^{k-1} / k] = 1 - \frac{1}{2} + \frac{1}{3} - \ldots \qquad ...(1)$$

This is alternating harmonic series, which is known to be conditionally convergent. If $\mathrm{E}(X)$ exists, it should have a unique value. However, examine (1) by *rearranging* its terms. Let it be $\mathrm{E}_1(X)$; then

$$\mathsf{E}_1(X) = 1 - \frac{1}{2} - \frac{1}{4} + \frac{1}{3} - \frac{1}{6} - \frac{1}{8} + \frac{1}{5} - \frac{1}{10} - \frac{1}{12} + \frac{1}{7} + \dots$$

$$= \left(1 - \frac{1}{2}\right) - \frac{1}{4} + \left(\frac{1}{3} - \frac{1}{6}\right) - \frac{1}{8} + \left(\frac{1}{5} - \frac{1}{10}\right) - \frac{1}{12} + \dots$$

$$= \frac{1}{2} - \frac{1}{4} + \frac{1}{6} - \frac{1}{8} + \dots = \frac{1}{2}\left(1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \dots\right) = \frac{1}{2}\mathsf{E}(X).$$

Thus, $\mathsf{E}_1(X) \neq \mathsf{E}(X)$. This is the reason that insist on absolute convergence for the definition of $\mathsf{E}(X)$, so as to get its unique vlaue.

Observation. Let $h(X)$ be a function of the random variable X and set $Y = h(X)$. To find $\mathsf{E}[h(X)]$, i.e. $\mathsf{E}(Y)$ we require the distribution function of Y; so

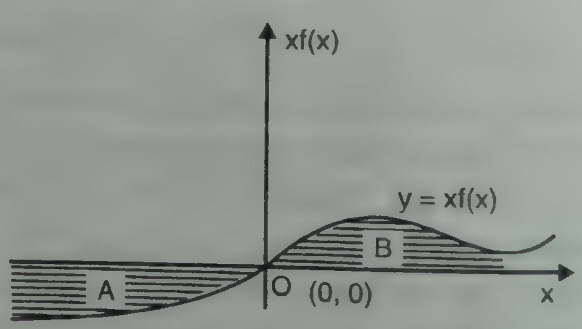
$$\mathsf{E}(Y) = \Sigma y_i P(Y = y), \text{ if } Y \text{ is discrete; } \quad \mathsf{E}(Y) = \int_R y f_Y(y)\, dy \text{ if } Y \text{ is continuous}$$

where we assume the knowledge of the distribution of Y.

As a matter of fact, there is no need to derive the distribution of Y and we show in Art. 5-12 (LUS), that

$$\mathsf{E}[h(X)] = \Sigma h(x_i) P(X = x_i), \ \forall\, x_i \text{ if } X \text{ is discrete; } \quad \mathsf{E}[h(X)] = \int_R h(x)\, f_X(x)\, dx, \text{ if } X \text{ is continuous.}$$

Note. On intutive grounds, we explain why do we require the absolute convergence of the series in (1) or integral in (2). Suppose X is continuous with p.d.f. $f(x) \geq 0$. Then $xf(x) \geq 0$, if $x \geq 0$ and $xf(x) \leq 0$ if $x < 0$. The graph of $y = xf(x)$ might be as shown in the adjoining figure. Let A denote the area which lies above the curve $y = xf(x)$ and is bounded above by the x-axis and let B denote the area which lies under the curve $y = xf(x)$ and is bounded below by x-axis. Then, by the geometric meaning of definite integral,



$$\text{(i) } \int_{-\infty}^{\infty} xf(x)\, dx = B - A, \text{ and (ii) } \int_{-\infty}^{\infty} |\, x \,|\, f(x)\, dx = B + A.$$

There are four possibilities :

(a) $B = \infty$, $A = \infty$,

(b) $B = \infty$, $0 \leq A < \infty$,

(c) $0 \leq B < \infty$, $A = \infty$

(d) $0 \leq B < \infty$, $0 \leq \infty A < \infty$.

In cases (a), (b) and (c), we see that $B + A = \infty$ and $B - A$ is not a finite number ($\infty - \infty$ is indeterminate). That is, integral (ii) is ∞ and integral (i) does not exist. In case (d), $A + B$ and $B - A$ are both finite. Hence, integral (i) is defined whenever integral (ii) is finite ($<\infty$).

5-11. Geometric Formulation of Expectation

If F denotes the distribution function of a continuous r.v. X, then

$$\boxed{E(X) = \int_0^{\infty} [1 - F(x)]\, dx - \int_{-\infty}^{0} F(x)\, dx = \int_0^{\infty} \bar{F}(x)\, dx - \int_{-\infty}^{0} F(x)\, dx}$$

provided the integrals exist finitely.

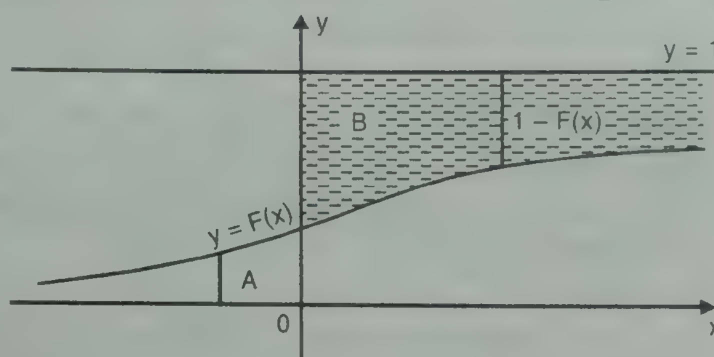
Proof. Let f be the density funciton of r.v. X. We may interpret $f(x)\, dx = dF(x)$. Now

$$E(X) = \int_{-\infty}^{\infty} xf(x)\, dx = \int_{-\infty}^{0} xd\, F(x) + \int_0^{\infty} xd\, F(x) = \int_{-\infty}^{0} xd\, F(x) - \int_0^{\infty} xd\, [1 - F(x)]$$

$$= [xF(x)]_{-\infty}^{0} - \int_{-\infty}^{0} F(x)\, dx - \{x[1 - F(x)]\}_0^{\infty} + \int_0^{\infty} [1 - F(x)]\, dx. \qquad \text{[Integ. by parts]}$$

Now $\lim\limits_{x \to \infty} x[1 - F(x)] = \lim\limits_{x \to \infty} x \int_x^{\infty} dF(t)\, dt \le \lim\limits_{n \to \infty} \int_x^{\infty} tdF(t) = 0.$ $\qquad \left[\int_{\infty}^{\infty} \equiv 0\right]$

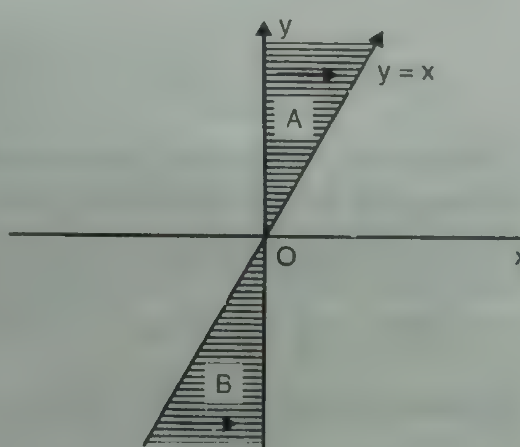
The integral above tends to zero as $x \to \infty$, because it is the tail of a convergent integral. Simiarly, $\lim\limits_{x \to -\infty} x\, F(x) = 0$. Hence the above result reduces to



$$E(X) = \int_0^{\infty} [1 - F(x)]\, dx - \int_{-\infty}^{0} F(x)\, dx. \qquad ...(1)$$

Note. The geometric content of Eq. (1) is $\mu = E(X) =$ Area B − Area A.

Also $\boxed{E(X) = \int_0^{\infty} P(X > t)\, dt - \int_{-\infty}^{0} P(X \le t)\, dt}$ $\qquad ...(2)$



For the integer-valued r.v. X, Eq. (1) becomes

$$\boxed{E(X) = \sum_{k=1}^{\infty} [1 - F_X(k)] - \sum_{k=0}^{\infty} F_X(-k)} \qquad ...(3)$$

An Alternative Proof. Let $f(x)$ be the p.d.f. of X.

$$\int_0^{\infty} [1 - F(x)]\, dx = \int_0^{\infty} P(X > x)\, dx = \int_0^{\infty} \left(\int_x^{\infty} f(y)\, dy\right) dx$$

$$= \int_0^{\infty} f(y) \left(\int_0^{y} dx\right) dy \text{ (Interchange limits in region } A)$$

$$= \int_0^{\infty} yf(y)\, dy.$$

$$\int_{-\infty}^{0} F(x)\, dx = \int_{-\infty}^{0} P(X \le x)\, dx = \int_{-\infty}^{0} \left(\int_{-\infty}^{x} f(y)\, dy\right) dx$$

$$= \int_{-\infty}^{0} f(y) \left(\int_y^{0} dx\right) = -\int_{-\infty}^{0} yf(y)\, dy \text{ (change of order in region } B)$$

$$E(X) = \int_{-\infty}^{\infty} tf(t)\, dt = \int_0^{\infty} yf(y)\, dy + \int_{-\infty}^{0} yf(y)\, dy = \int_0^{\infty} [1 - F(x)]\, dx - \int_{-\infty}^{0} F(x)\, dx.$$

***Cor.* 1.** If X is *non-negative* r.v. with c.d.f. 'F', then, $\mathsf{E}(X) = \int_0^\infty [1 - F(x)]\, dx = \int_0^\infty P(X > x)\, dx.$

***Cor.* 2.** If X is non-negative *integer-valued* variate, then $\mathsf{E}(X) = \sum_{j=1}^{\infty} P(X \geq j).$

A direct proof of Cor. 2 is much more illuminating. [See Ex. 5.4(a)]

Note. For a variate, which is partly continuous and partly discrete, we use Art. 5-11(1), where geometric interpretation of $\mathsf{E}(X)$ is given.

For illustrations of Cor. 1 and Cor. 2, we may find $\mathsf{E}(X)$ for expo (λ) and gem (p).

5-20. *A law of Unconscious Statistician (L.U.S.)*

Let X be a r.v. and h be any real-valued function of X. Then

$$\mathsf{E}[h(X)] = \Sigma_x h(x)\, p(x),\ [p(x) > 0 \textit{ is p.m.f. of discrete } \text{r.v. } X] \qquad \ldots(1)$$

$$= \int_{-\infty}^{\infty} h(x)\, f(x)\, dx\,,\ [f(x) \text{ is p.d.f. of continuous r.v. } X] \qquad \ldots(2)$$

We unconsciously use density of X instead of the density of $Y = h(X)$.

Proof. (1) Let X be discrete and suppose X assumes a finite number of values. Let h_i, $1 \leq i < n$ be n possible values of $h(X)$. For reach i, $1 \leq i \leq n$, let x_{ij} denote the values of X such that $h(x_{ij}) = h_i$, $1 \leq j \leq m$. Then

$$P\{h(X) = h_i\} = \sum_{j=1}^{m} P\{X = x_{ij}\} = \sum_{j=1}^{m} f(x_{ij}).$$

$$\therefore \quad \mathsf{E}[h(X)] = \sum_{i=1}^{n} h_i P[h(X) = h_i] = \sum_{i=1}^{n} h(x_{ij}) \sum_{j=1}^{m} f(x_{ij}) = \sum_{i=1}^{n}\sum_{j=1}^{m} h(x_{ij})\, f(x_{ij}) = \Sigma_x h(x)\, f(x). \qquad \ldots(i)$$

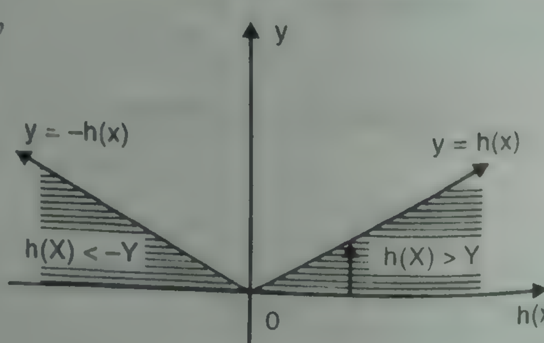
If X assumes countably infinite values with positive probability, properties of absolutely convergent series allow the same conclusion.

(2) Let X be continuous. The geometric content of $\mathsf{E}(Y)$ is

$$\mathsf{E}(Y) = \int_0^\infty [1 - F(y)]\, dy - \int_{-\infty}^{0} F(y)\, dy = \int_0^\infty P(Y > y)\, dy - \int_0^\infty P(Y < y)\, ay \qquad \ldots(ii)$$

Replace Y by $h(X)$; denote the sets : $A_y = \{x : h(x) > y\}, B_y = \{x : h(x) < -y\}$; then (ii) provides

$$\begin{aligned}
\mathsf{E}[h(X)] &= \int_0^\infty P\{h(X) > y\}\, dy - \int_0^\infty P\{h(X) < -y\}\, dy \\
&= \int_0^\infty dy \int_{A_y} f(x)\, dx - \int_0^\infty dy \int_{B_y} f(x)\, dx \\
&= \int_{A_0} f(x)\, dx \int_0^{h(x)} dy - \int_{B_0} f(x)\, dx \int_0^{h(x)} dy \\
&= \int_{A_0} f(x)\, dx \int_0^{h(x)} dy - \int_{B_0} f(x)\, dx \int_0^{h(x)} dy \\
&= \int_{A_0} h(x)\, f(x)\, dx + \int_{B_0} h(x)\, f(x)\, dx \\
&= \int_{-\infty} h(x)\, f(x)\, dx \qquad [A_0 \cup B_0 = R = \{-\infty < x < \infty\}]
\end{aligned}$$



(Interchange limits)

5-21. Illustrations of the L.U.S. : Two-way Expectation

***Qn.* 1.** Let $Y = \sin X$ where $X \sim U(0, \pi/2)$, i.e. $f(x) = 2/\pi$, $0 < x < \pi/2$. Compute $E(Y)$ by 2 ways.

Solution. (i) $E(Y) = E(\sin X) = \int_0^{\pi/2} \frac{2}{\pi} \sin x \, dx = \frac{1}{\pi}$. [L.U.S. method]

(ii) $F_Y(y) = P(Y \le y) = P(\sin X \le y) = P(X \le \sin^{-1} y) = \int_0^{\sin^{-1} y} \frac{2}{\pi} dx = \frac{2}{\pi} \sin^{-1}(y)$.

$\therefore$ $f_Y(y) = F'_Y(y) = 2/\pi\sqrt{1-y^2}$, $0 < y < 1$.

$$E(Y) = \int_0 y \, . \, f_Y(y) \, dy = \frac{2}{\pi}\int_0^1 \frac{y \, dy}{\sqrt{1-y^2}} = \frac{2}{\pi}\int_0^{\pi/2} \sin\theta \, d\theta = \frac{2}{\pi}. \qquad [y = \sin\theta]$$

Thus, $E(Y)$ evaluted by two different ways has the same result.

***Qn.* 2.** Let $Y = e^x$ where $X \sim U(0, 1)$, i.e. $f(x) = 1$, $0 < x < 1$. Compute $E(Y)$ by two ways.

Solution. (i) $E(Y) = E(e^x) = \int_0^1 1 \, . \, e^x \, dx = e = 1$. [L.U.S. method]

(ii) $F_y(y) = P(Y \le y) = P(e^X \le y) = P(X \le \log y) = \int_0^{\log y} 1 \, dx = \log y$.

$\therefore$ $f_Y(y) = F'_Y(y) = (1/y)$ $1 \le y \le e$

$\therefore$ $E(Y) = \int_1^e f_Y(y) \, . \, y \, dy = \int_1^e dy = e - 1$.

Thus, $E(Y)$ evaluted by two different ways has the same result.

***Qn.* 3.** Let $Y = X^2$ where $X \sim$ bin $(2, 1/2)$. Compute $E(Y)$ by two ways.

Solution. $E(Y) = E(X^2) = 0^2 \, . \, \frac{1}{4} + 1^2 \, . \, \frac{1}{2} + 2^2 \, . \, \frac{1}{4} = \frac{3}{2}$. $\left[P(X = k) = \binom{2}{k}\frac{1}{4}\right]$

Let $Y = X^2$, then $g(0) = 1/4$, $g(1) = 1/2$, $g(4) = 1/4$ and $E(Y) = 0 \, . \, (1/4) + 1 \, . \, (1/2) + 4(1/4) = 3/2$.

Theorem. If F is the c.d.f. of a r.v. X and $E(X^n)$ exists, then for any a and $1 \le k \le n$,

$$\boxed{E(X-a)^k = k\int_a^{\infty}(x-a)^{k-1}[1-F(x)] \, dx - k\int_{-\infty}^{a}(x-a)^{k-1}F(x) \, dx} \qquad ...(1)$$

Proof. Let f be the density function of X and interpret $f(x) \, dx = dF(x)$. Now

$$E(X-a)^k = \int_{-\infty}^{\infty}(x-a)^k f(x) \, dx = \int_{-\infty}^{a}(x-a)^k \, dF(x) - \int_a^{\infty}(x-a)^k \, d[1-F(x)]$$

$$= [(x-a)^k F(x)]_{-\infty}^{a} - \int_{-\infty}^{a} F(x) \, d(x-a)^k - \left\{(x-a)^k[1-F(x)]\right\}_a^{\infty} + \int_a^{\infty}[1-F(x)] \, d(x-a)^k$$

Now $\lim\limits_{x\to\infty}(x-a)^k[1-F(x)] = \lim\limits_{x\to\infty}(x-a)^k\int_x^{\infty} dF(t) \le \lim\limits_{x\to\infty}\int_x^{\infty}(t-a)^k \, dF(t) = 0$

The last integral tends to zero as $x \to \infty$, because it is the tail of a convergent integral. Similarly, $(x-a)^k F(x) \to 0$ as $x \to -\infty$. Hence the above result reduces to

$$E(X-a)^k = k\int_a^{\infty}(x-a)^{k-1}[1-F(x)]\,dx - k\int_{-\infty}^{a}(x-a)^{k-1}F(x)\,dx.$$

Cor. $E(X^k) = k\left\{\int_a^{\infty} x^{k-1}[1-F(x)]\,dx - \int_{-\infty}^{a} x^{k-1}F(x)\,dx\right\}$ ...(i)

$$\boxed{E(X)^k = k\int_0^{\infty} x^{k-1}[1-F_X(x)+(-1)^k F_X(-x)]\,dx} \quad \text{...(ii)}$$

Note that (ii) follows from (i) by using $x = -t$ in 2nd integral in (i) and then changing dummy t to dummy x.

5-30. *Some Important Properties of Expectation*

Let X, Y be random variables and a, b, c, be constants. Then for regular functins g, h,

(i) $E(c) = c$ [Constant-invariance]

(ii) $E(aX + b) = aE(X) + b$

(iii) $P(X = c) = 1 \Rightarrow E(X) = c$ [Degenerate variate]

(iv) $P\{a < X \le b\} = 1 \Rightarrow a < E(X) \le b$ [Certainty Limits]

(v) $E\{ag(X) + bh(X)\} = aE\{g(X)\} + bE\{h(X)\}$ [Linearity]

(vi) $E(X) = \sum_{n=1}^{\infty} P(X \ge n) = \sum_{n=0}^{\infty} P(X > n).$ [X is non-neg integer-valued r.v.]

(vii) $E(X-a)^k = k\int_a^{\infty}(x-a)^k\,[1-F(x)]\,dx - k\int_{-\infty}^{a}(x-a)^{k-1}F(x)\,dx$

(viii) $E\{ag(X)+bh(Y)+c\} = aE[g(X)] + bE[h(Y)] + c,$ [Linearity (§5-51)]

The linear property of E shall be coded : Lin E.

Proofs of Some Operational Properties of E

(i) Treating $X \equiv c$ as a Degenerate r.v. that assumes its only possible value c with probability 1, we get

$$E(X) = c\,.\,P(X = c) = c\,.\,1 = c. \qquad [X(\omega) = c,\ \forall\ \omega \in \Omega]$$

(ii) Suppose X is continuous with p.d.f. $f_X(x)$, then

$$E[ag(X)+bh(X)] = \int_{-\infty}^{\infty}[ag(X)+bh(X)]\,f_X(x)\,dx \qquad \text{[by Def.]}$$

$$= a\int_{-\infty}^{\infty} g(X) f_X(x)\,dx + b\int_{-\infty}^{\infty} h(x)\,f_X(x)\,dx = aE[g(X)] + bE[h(X)] \quad \text{...(i)}$$

If $b = 0$, this result gives $E[bg(X)] = bE[g(X)]$

Cor. $E(aX + b) = aE(X) + b.$

Replace $g(X)$ by X and $h(X)$ by 1 in (i) to get the result.

More Properties

1. If $X \geq 0$, then $E(X) \geq 0$.

Proof. Assume X is absolutely continuous. Since X is non-negative, $f(x) = 0$ for $x < 0$. Hence, provided $E(X)$ exists,

$$E(X) = \int_{-\infty}^{\infty} xf(x)\,dx = \int_0^{\infty} xf(x)\,dx > 0$$

2. The expected value of a *bounded* r.v. X always exists.

Proof. Assume X is absolutely continous and bounded, i.e. $|X| \leq B$, so that $f(x) = 0$, for $x > B$. Now

$$\int_{-\infty}^{\infty} |x| f(x)\,dx = \int_{-B}^{B} |x| f(x)\,dx \leq B\int_{-B}^{B} f(x)\,dx = B.$$

If X is discrete and bounded, i.e.$|X_i| \leq B$, $\forall i$, so that $p_i = 0$, for $|x_i| > B$, then, for all i

$$\Sigma |x_i| p_i \leq B(\Sigma p_i) = B.$$

Thus, in either case, $E(|X_i|) \leq B$, whence expectation necessary exists.

3. ***Relation between E and P.*** *Expected value of an indicator variate* $E(\mathbb{I}_A) = P(A)$

Expected value of an indicator variate $X \equiv I_A$ is given by

$$E(I_A) = 1 . P(I_A = 1) + 0 . P(I_A = 0) = P(I_A = 1) = P(A) = \int_A dF_X(x)$$

This result is valuable, since it is frequently used to find $P(A)$ rather than to evaluate $E(X)$.

5-31. *Worked-out Problems*

***Example* 1.** If the possible values of a r.v. X are 0, 1, 2, 3, ..., then

(i) $E(X) = \Sigma P(X > n)$, $n = 0, 1, 2, 3, ...$ ...(1)

(ii) $\sum_{k=0}^{\infty} kP(X > k) = \frac{1}{2}[E(X^2) - E(X)]$. ...(2)

Solution. (i) We assume that the series $\Sigma P(X > x)$ is convergent. Now, using $P(X = k) = p_k$, we observe that

$$P(X > 0) = P\{(X = 1) \uplus (X = 2) \uplus (X = 3) \uplus ...\} = p_1 + p_2 + p_3 + ...$$
$$P(X > 1) = P\{(X = 2) \uplus (X = 3) \uplus (X = 4) \uplus ...\} = p_2 + p_3 + p_4 + ...$$
$$P(X > 2) = p_3 + p_4 + p_5 + ..., \quad P(X > k) = p_{k+1} + p_{k+2} + ...$$

and so on. Adding all these terms, we get

$$\sum_{n=0}^{\infty} P(X > n) = P(X > 0) + P(X > 1) + P(X > 2) + ... = p_1 + 2p_2 + 3p_3 + 4p_4 + ... = \sum_{k=0}^{\infty} kp_k = E(X).$$

Note. If $E(X)$ fails to exist, then both sides of (1) are equal to ∞.

(ii) Since $k = 0$ yields zero term, we commence with $k = 1$. We may also put $P(X = j) = p_j$, so

$$\sum_{k=1}^{\infty} kP(X > k) = \sum_{k=1}^{\infty} k \sum_{j=k+1}^{\infty} p_j = \sum_{j\geq 1} p_j \sum_{k=0}^{j-1} k = \sum_{j\geq 1} \frac{1}{2} j(j-1)\,p_j = \frac{1}{2}\left[\sum_{j\geq 1} j^2 p_j - \sum_{j\geq 1} jp_j\right]$$
$$= \tfrac{1}{2}[E(X^2) - E(X)].$$

Example 2. Gambling system (Bold Play) : A fair coin is tossed. If 'h' occurs, you win Rs 1. If 't' occurs, you lose Rs 1. If you win, you quit. But if you lose, you double your bet so that on the second toss you either win Rs. 2 or lose Rs. 2. You continue playing in this manner, quiting the moment you win and doubling your bet if you lose. Since your bank-deposits is finite, you can afford at most n attempts. If X denotes your winnings (or losings), evaluate $E(X)$.

Solution.

$$\Omega = \{h_1, t_1h_2\ t_1t_2h_3, \ldots, t_1t_2, \ldots t_{n-1}h_n, t_1t_2, \ldots t_n\}$$

$$X(h_1) = X(t_1h_2) = X(t_1t_2h_3) = \ldots = X(t_1t_2 \ldots t_{n-1}h_n) = 1 \quad \text{(gain)}$$

$$X(t_1t_2 \ldots t_n) = 2^0 + 2^1 + \ldots + 2^{n-1} = 2^{n-1} \text{ (Loss)}$$

$X(t_1\ t_2 \ldots t_kh_{k+1}) = (2^0 + 2^1 + \ldots + 2^{k-1})$ loss $+ 2^k$ (gain) $= -(2^k - 1) + 2^k = 1$ [general term]

$f(x) =$ Prob. of losing $= P\{t_1 \ldots t_n\} = 1/2^n$. Prob. of winning $= 1 - 1/2^n$.

$\therefore f_X(1) = 1 - (1/2)^n$; $f_X(-2^n + 1) = (1/2)^n$. (p.m.f.) $E(X) = 1[1 - (1/2)^n)] + (-2^n + 1)(1/2)^n = 0$.

Note. $\lim_{n\to\infty} f_X(1) = 1$, so winning Rs. 1 is almost certain.

Problems with Solutions Provided at the End of the Text

1*. Prove that the mean is not defined for each of the following random variables :

(a) $f(x) = 1/x,\ x = 2^n,\ n = 1, 2, 3, \ldots$

(b) $f(x_i) = 2/(3)^i,\ x_i = (-1)^{i+1}3^i/i,\ i = 1, 2, \ldots$

(c) $f(x) = 1/x(x + 1),\ x = 1, 2, \ldots;\ f(x) = 0$, elsewhere

(d) $f(x) = 1/(x + 1)^2,\ 0 < x < \infty$.

(e) $f(x) = 1/\pi(1 + x^2),\ -\infty < x < \infty$ [Cauchy density]

2*. Show that, for a random variable X, $E(X)$ exists if $E(|X|)$ exists and in this case $|E(X)| \le E(|X|)$.

3*. Let $X_1, X_1, \ldots, X_n$ be *i.i.d.* variates and put $S_n = X_1 + X_2 + \ldots X_n$.

Find (i) $E(S_k / S_n)$, (ii) $E(X_1 | S_n = t)$, $(k \le n)$.

4*. If X has p.d.f. $f(x) = e^{-x},\ x > 0$, find $E(Y)$, where

$$Y = X, \text{ if } X > 3;\ Y = 2X + 3, \text{ if } X \le 3.$$

5*. If $F(x)$ is the c.d.f. of a continuous variate X defined for $0 < x < b < \infty$, prove that

$$E(X) = \int_0^b [1 - F(x)\,dx]$$

6*. Let $X_1, X_2, \ldots, X_m$ be i.i.d. integer-valued variates with $P(X_1 = k) = p_k$.

Let $r_n = \sum_{k=n}^{\infty} p_k$ and $Y = \min\{X_1, \ldots X_m\}$. Show that $E(Y) = \sum_{n=1}^{\infty}(r_n)^m$. Evaluate it when $X_j \sim$ gem (p).

7*. A jar has n chips numbered $1, 2, \ldots, n$. A person draws a chip, returns it, draws another, returns it and so on, until a chip is drawn that has been drawn before. Let X be the number of drawings. Find $E(X)$.

8*. If a r.v. X has possible values $x_1, x_2, \ldots, x_n$ where $x_i < x_{i+1}$, $1 \le i \le n-1$, then
$\mathsf{E}(X) = \Sigma(x_i - x_{i-1})\, P(X \ge x_i)$, $(x_0 \equiv 0)$

Another variate Y has the same set of values as X and additionally satisfies
$P(X \ge x) \ge P(Y \ge x)$, for all x ...(1)
Show that $\mathsf{E}(X) \ge \mathsf{E}(Y)$.

9*. Let $a_1, a_2, \ldots, a_n$ be arbitrary real numbers and $A_1, A_2, \ldots, A_n$ be events. Show that

$$\sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j\, P(A_i A_j) \ge 0.$$

10*. Variate X has p.d.f. 'f' : $f(x) > 0$ at $x = -1, 0, 1$ and $f(x) = 0$, elsewhere. (i) If $f(0) = \frac{1}{2}$, find $\mathsf{E}(X^2)$. (ii) If $f(0) = \frac{1}{2}$, $\mathsf{E}(X) = 1/6$, find $f(-1)$ and $f(1)$.

11*. What is the mean number of face cards obtained in a Bridge-hand from a well-shuffled deck ?

12*. A coin is tossed until a head appears. What is the expectation of the number of tosses ?

13*. What is the expectation of the number of failures preceding the first success, in an indefinite series of independent trials, with constant probability p of success in each trial ?

14*. An urn contains n cards numbered from 1 to n. Two cards are drawn at a time. Find the mathematical expectation of the product of the numbers on the cards.

15*. Find the expected value of $Y = (X-1)^2$, when variate X has the *p.m.f.* give by

X :	0	1	2	3
$p(x)$:	1/3	1/2	1/24	1/8

16*. If X_1, X_2, X_3 are i.i.d. $U(0, a)$ distributed, find $\mathsf{E}[1-(X/a)]^2$ where $X = \min(X_1, X_2, X_3)$.

5-40. *Expectation of a Variate Conditioned by an Event*

Let X be a variate on a given probability space $(\Omega, \mathcal{F}, P)$ and A an event with $P(A) > 0$. The conditional expectation of X, given that the event A has occurred, is denoted by $\mathsf{E}(X \mid A)$, and is defined by

$$\mathsf{E}(X \mid A) = \mathsf{E}(XI_A)/P(A), \quad [P(A) > 0] \qquad \text{...(1)}$$

Note that, $\mathsf{E}(XI_A) = \Sigma x_i P(X = x_i)$, $\forall\, x_i \in A$. So (1) can be expressed as

$$\mathsf{E}(X \mid A) = \Sigma x_i\, P(x_i)/P(A),\ x_i \in A \qquad \text{...(2)}$$

5-41. *Multi-Stage Rule for Expectation (Multistage E-Rule)*

If A and B are possible events and $[P(A \cup B) > 0]$ then

$$\mathsf{E}(X \mid A \cup B) = \frac{P(A)\mathsf{E}(X \mid A) + P(B)\mathsf{E}(X \mid B) - P(AB)\mathsf{E}(X \mid AB)}{P(A \cup B)} \qquad \text{...(1)}$$

Proof. Write $C_x = \{X = x\}$ and $P\{C_x \mid H\} - f(x \mid H)$. Now

$$f(x/A \cup B) = \frac{P\{C_x \cap (A \cup B)\}}{P(A \cup B)} = \frac{P\{C_x A \cup C_x B\}}{P(A \cup B)}$$

$$= \frac{[P(AC_x) + P(BC_x) - P(ABC_x)]}{P(A \cup B)}$$

$$= \frac{[P(A)\,f(x \mid A) + P(B)\,f(x \mid B) - P(AB)\,f(x \mid AB)]}{P(A \cup B)}$$

$\mathsf{E}(X \mid A \cup B) = \Sigma x f(x \mid A \cup B)$ (over range of x-values)

$$= \frac{P(A)\,\Sigma x f(x \mid A) + P(B)\,\Sigma x f(x \mid B) - P(AB)\,\Sigma x(x \mid AB)}{P(A \cup B)}$$

$$= \frac{P(A)\,\mathsf{E}(X \mid A) + P(B)\,\mathsf{E}(X \mid B) - P(AB)\,\mathsf{E}(x \mid AB)}{P(A \cup B)}$$

***Cor.* 1.** If A and B are disjoint, $P(AB) = 0$, and we obtain

$$\mathsf{E}(X \mid A \uplus B) = [P(A)\,\mathsf{E}(X \mid A) + P(B)\,\mathsf{E}(X \mid B)/P(A \uplus B).$$

Cor.* 2. *Total-*E *Rule. If $A_1, A_2, \ldots, A_n, \ldots$ constitute a *Partition* of Ω, $[\uplus A_i = \cap]$ then

$$\mathsf{E}\left[X \biguplus_{i=1}^{\infty} A_i\right] = \sum_{i=1}^{\infty} P(A_i)\,\mathsf{E}(X \mid A_i) \Rightarrow \mathsf{E}(X) = \Sigma P(A_i)\,\mathsf{E}(X \mid A_i).\ [X \mid \cap \equiv X]$$

Cor.* 3. *Multistage* E*-Rule

For a r.v. X associated with occurrence of event A or $\bar{A}$.

$$\mathsf{E}(X) = P(A) + \mathsf{E}(X \mid A) + P(\bar{A})\,\mathsf{E}(X \mid \bar{A})$$

Proof. Write $C_X = \{X = x\}$, $P(C_X) = P(X = x) = f(x)$ say. Then

$$f(x) = P\{C_x \cap (A \uplus \bar{A})\} = P(AC_x \uplus \bar{A}C_x) = P(AC_x) + P(\bar{A}C_x)$$

$$= P(A)\,P|C_x \mid A| + P(\bar{A})\,P|C_x \mid \bar{A}) \qquad [P(C_x \mid A) = f(x \mid A)]$$

$\therefore$
$$\mathsf{E}(X) = \Sigma x f(x) = P(A)\,\Sigma x f(x \mid A) + P(\bar{A})\,\Sigma x\,.\,f(x \mid \bar{A})$$

$$= P(A)\,\mathsf{E}(X \mid A) + P(\bar{A})\,\mathsf{E}(X \mid \bar{A})$$

Extension. If X is associated with occurrence of events $A_1, \ldots, A_n$ where $\uplus A_j = S_j$ then

$$\mathsf{E}(X) = P(A_1)\,\mathsf{E}(X \mid A_1) + P(A_2)\,\mathsf{E}(X \mid A_2) + \ldots + P(A_n)\,\mathsf{E}(X \mid A_n)$$

5-42. *Worked-out Problems*

***Example* 1.** A man wins a rupee for head and loses a rupee for tail when a coin is tossed. Suppose that he tosses once and quits if he wins; but tries once more if he loses on the first toss. What are his expected winnings ?

Solution. We shall use Multi-Stage Rule; define events : $A = \{H_1\}$, = (head on first toss); $B = \{T_1 H_2\}$, = (Tail on first toss, Head on 2nd toss); and $C = \{T_1 T_2\}$ = {Tails on both tosses}. If X denotes his random earning, then by Total-E rule

$$\mathsf{E}(X) = P(A)\,\mathsf{E}(X \mid A) + P(B)\,\mathsf{E}(X \mid B) + P(C)\,\mathsf{E}(X \mid C) \qquad \ldots(1)$$

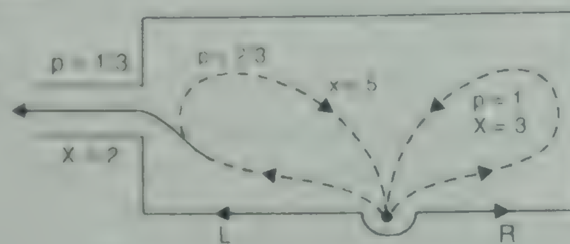
Assuming coin in unbiased, we have : $P(A) = \frac{1}{2}$, $P(B) = \frac{1}{4}$, $P(C) = \frac{1}{4}$;

$E(X \mid A) = +1$, $E(X \mid B) = -1 + 1 = 0$, $E(X \mid C) = -1 - 1 = -2$.

$E(X) = \frac{1}{2} \cdot 1 + \frac{1}{4} \cdot 0 + \frac{1}{4}(-2) = 0$

Example 2. A rat is trapped in a maze. Initially he has to choose one of two directions. If he goes to the right, then he will wander around in the maze for 3 minutes and will then return to his initial position. If he goes to the left, then with probability $\frac{1}{3}$, he will depart the maze after 2 minutes of travelling and with probability $\frac{2}{3}$, he will return to his initial position after 5 minutes of travelling. Assuming that the rat is, at all times, equally likely to go to the left or the right, find the expectation and the variance of the amount of time that he will be trapped in the maze.

Solution. **The geography of the problem is indicated in the figure. Let X denote the number of minutes in the maze. We use mgf method and utilize multi-stage E-rule.**



$$E(e^{tX}) = E\{e^{tX} \mid L\}\,P(L) + E\{e^{tX} \mid R\}\,P(R)$$

As $P(L) = P(R) = \frac{1}{2}$, the above turns out to be

$$2M(t) = E(e^{tX} \mid L) + E(e^{tX} \mid R) \qquad \text{...(1)}$$

Now

$$E(e^{tX} \mid L) = \tfrac{1}{3}e^{2t} + \tfrac{2}{3}E[e^{t(X+5)}] = \tfrac{1}{3}e^{2t} + \tfrac{2}{3}\{e^{5t}E(e^{tX})\}$$

$$E(e^{tX} \mid R) = 2\,.\,E[e^{t(X+3)}] = e^{3t}E(e^{tX}).$$

Explanation. Turning to right (say), the rat spends 3 minutes and reaches the initial position, from where he has to start afresh, again with X minutes further captivity. Similar explanation holds for $E(e^{tX} \mid L)$.

We substitute the preceding two results into (1), collect the terms to get

$$M(t) = e^{2t}/(6 - 2e^{5t} - 3e^{3t}) \qquad \text{...(2)}$$

$$= e^{2t}/[1 - 19t - (77/2)t^2 \ldots]$$

$$\ln M(t) = 2t - \ln\{1 - [19t + (77/2)t^2 + \ldots]\}, \text{ use expansion.}$$

$$K(t) = \ln M(t) = 2t + [19t + (77/2)t^2 + \ldots] + \tfrac{1}{2}[19t + (77/2)t^2 + \ldots]^2 + \ldots$$

$$E(X) = K_1 = 21,\ \mathrm{Var}(X) = K_2 = 438$$

Problems with Solutions Provided at the End of the Text

1*. The average height of the boys in a class is 65″. The average height of the girls is 60″. Sixty percent of the class is male. What is the average height of a student, in the class ?

2*. A whole number X is chosen at random between numbers 1 and 5 inclusive. The number Y is then chosen at random between numbers 1 and X. Find the expected value of Y.

3*. A card is picked at random from among n cards numbered 1, 2, ..., n. If the card is k, then a second card is picked from among cards 1, 2, ... k. Denoting the first number selected by Y and the second by X, show that $E(X) = (n + 3)/4$.

Exercise 5(a)

1. (a) If $P(X = x) = 1/5, x = 1, 2, ..., 5$; find $E(X + 2)^2$.

   (b) A variate X has the probability distribution : $f(-3) = 1/6, f(6) = 1/2, f(9) = 1/3$. Find $E(X)$, $E(X^2)$ and using the law of Expectation, show that $E(2X + 1)^2 = 209$.

2. (a) n people, including A and B, stand up in a random order in a (i) row, (ii) ring. Find the expected number of people between A and B.

   (b) A contractor is to choose between two jobs. The first promises a profit of Rs. 1,20,000 with a probability of $\frac{3}{4}$ or a loss of Rs. 30,000 due to delays with a probability of 1/4. The second job premises a profit of Rs. 1,80,000 with a probability of $\frac{1}{2}$ or a loss of Rs. 45,000 with a probability of $\frac{1}{2}$. Which job should the contractor choose so as to maximize his expected profit ? **[Ans.** Job 1 is preferable]

3. A and B throw one die for a prize of Rs. 44, which is to be won by the player who first throws a 6. If A has the first throw, show that their respective expectations are 24 and 20 units.

4. A bag contains 2 white and 3 black balls. Four persons A, B, C, D in the order named, each draws one ball without replacement. The first to draw a white ball receives Rs. 10. What are their expectations ? **[Ans.** 4, 3, 2, 1]

5. A man draws 2 balls from a bag containing 3 white and 5 black balls. If he is to receive Rs. 10, for every white ball which he draws and pay Rs. 5 for every black one, show that his expectation is 5/4.

6. A lot of 10 items is totally rejected if two items chosen at random from the lot contain 1 or more defectives. Otherwise, it is accepted. Each item costs Rs. 700 and sold for Rs. 900. If the lot contains 2 defective items, what is the manufacturer's expected profit ? **[Ans.** Loss of Rs. 1400]

7. A player tosses 3 fair coins. He wins Rs. 8, if three heads occur; Rs. 3, if 2 heads occur and Rs. 1, if one head occurs. If the game is to be fair, how much should he lose if no heads occur ? **[Ans.** Rs. 20]

8. A makes a bet with B of Rs. 9 to Rs. 6 that in a single throw with 2 dice, he will throw 7 before B throws 4. Each has a pair of dice and they throw simultaneously until one of them wins, equal throws being disregarded. Find B's expectation. **[Ans.** Loss Re. 1]

9. Ajay plays two games, in each of which his probability of winning is p. If he loses, he loses his stake, but if he wins he gets double his stake. How should he divide his fortune between the two games to maximize his expected gains ?

   [Ans. Expected gain is indep. of the manner of subdivision of fortunes]

10. In a game of chance, a man is allowed to throw a coin indefinitely. He receives Rs. 1, 2, 3, ... if he throws a head at 1st, 2nd, 3rd, ... trials respectively. If the entry fee to participate in the game is Rs. 2/-, show that the expected value of the net gain is zero.

11. You play a game as follows : You pay Rs. 2 and toss a coin. If it lands heads, you are paid Rs. 11, otherwise you pay Rs. 4 more and roll a die. If the die shows 1, you are paid Rs. 12; if it does not, you pay Rs. 6 and draw a ball from an urn containing 2 black, 2 white and 3 red balls. If you draw a black or a red ball, you are paid Rs. 9, otherwise you pay Rs. 8. Find the expected pay off.

12. A bag contains n white and 2 black balls. Balls are drawn one by one without repalcement, until a black ball is drawn. If 0, 1, 2, 3, ..., white balls are drawn before the first black, a man is to receive $0^2, 1^2, 2^2, 3^2, \ldots$ rupees respectively. Prove that his expectation is $n(n + 1)/6$.

13. A person draws cards one by one from a pack until he draws all the aces. How many cards may he be expected to draw ?
[**Ans.** $E(X) = K\Sigma(x^4 - 6x^3 + 11x^2 - 64x)$, $K = 4/52.51.50.49$; $4 \le x \le 52$]

14. A game is placed as follows : A number is drawn at random from {1, 2, 3, 4, 5, 6, 7}. If the number is even, the player wins the number of rupees equal to the number drawn, otherwise, the player pays Rs. 3 and rolls two dice. If the sum is not 6, he pays Rs. 6; if the sum is equal to 6, he gets Rs. 12. What is the probability that he will make money if he plays once ? What is his expected pay off ?

15. An urn contains 100 tickets each bearing exactly one of the numbers from 1 to 100, no numbers being used more than once, A ticket (x) is drawn and its number noted. If $1 \le x \le 20$, the drawer loses Rs. 5. If $20 < x \le 35$, the drawer earns Rs. 3. If $35 < x < 95$, the drawer earns Rs. 10. If $x \ge 95$, the drawer earns the amount equal to that drawn by him. If Z represents the earning of the drawer in a single draw, show that $E(Z) = 97.62$.

16. Thirteen cards are drawn from a pack simultaneously. If the values of aces are 1, face cards 10 and others according to their denominations, show that the expectation of the total score in all the 13 cards is 85.

17. Two players A and B alternately roll a pair of dice. A wins if he gets 6 points before B gets 7 points and B wins, if he gets 7 points before A gets 6 points. If A takes the first turn, show that the probability that B wins is 31/61 and the expected number of trials for A to win is 6.

18. Show that the expected number of throws of a coin necessary to produce 3 heads is 6.

19. Let the p.d.f. of X be given by : $f(x) = 1/(x \log 3)$, $1 < x < 3$; $f(x) = 0$, elsewhere. Find $E(X)$, $E(X^n)$, Var (X), Var (X^n) for any positive integer n. [**Ans.** $E(X^k) = (3^k - 1)/k \log 3$, use $k = 1, 2, n, 2n$]

20. Let X be a continuous variate and F be its c.d.f.
Show that $E\{[F(X)]^n\} = (n+1)^{-1}$, and that $E[2F(X) - 1] = 0$.

21. Shwo that $E(X) = \pi^2/6$ when X is a random variable with p.m.f.
$P(X = n) = (2n+1)/[n(n+1)]^2$, $n = 1, 2, 3, \ldots$

22. The p.m.f. of a variate X is : $P(X = n) = 6/\pi^2 n^2$, $n = 1, -2, 3, -4, \ldots$ Does $E(X)$ exist ?

23. A discrete variate X assumes three values $-3, 0, 4$ and $P(X = 0) = \frac{1}{2}$ and $E(X) = 9/8$. Show that $P(X = -3) = 1/8$, and $P(X = 4) = 3/8$.

24. Let $f(x) = (1 + x^3)/6$, $0 \le x \le 2$; $f(x) = 0$, elsewhere. Let n be a non-negative integer. Show that
$$E(X^n \mid X \le 1) = (4/5)\,[(n+1)^{-1} + (n+4)^{-1}] \text{ and}$$
$$E(X^n \mid X > 1) = (4/9)\,[(2^{n+1} - 1)/(n+1)] - [(2^{n+4} - 1)/(n+4)]\}.$$

25. Discrete variates X and Y have the same finite set of possible values and $E(X) = E(Y) = m$. If for every possible value u, $P\{|X - m| \ge u\} \le P\{|Y - m| \ge u\}$, (i.e. X is more concentrated about m than Y], show that $\sigma_X^2 \le \sigma_Y^2$.

26. Let p, $(0 < p < 1)$ be the probability that a child is a son. Show that the expected number of sons in a family with n children, given that there is at least one son is $np/(1 - q^n)$.

27. A gamester has a disc with a freely revolving needle. The disc is divided into 100 equal sectors by thin lines and the sectors are marked 0, 1, 2, ..., 99. The gamester treats 5 or any multiple of 5 as lucky numbers and zero as a special lucky number. He allows the player to whirl the needle on a charge of Re 1. When the needle stops at the lucky number, the gamester pays back the player Rs. 2 and at special lucky number, he pays to the player Rs. 5. Is the game fair ? [**Ans.** Not fair; player loses Rs. 0.37 per game]

28. A bag contains $2n$ counters, of which half are marked with old numbers and a half with even numbers, the sum of all the numbers being S. A man is to draw two counters. If the sum of the numbers drawn is odd, he is to receive that number of rupees, if even, he is to pay that number of rupees. Show that his expectation is $S/[n(2n-1)]$ rupees.

29. A prisoner is trapped in a cell containing three doors. The first door leads to a tunnel which takes him to freedom after 2 hours travel. The second door leads to a tunnel which returns him to the cell after 3 hrs. travel. The third door leads to a tunnel which returns him to his cell after 4 hrs. Assuming that the prisoner is at all times equally likely to choose any one of the doors, what is the expected length of time until prisoner becomes free ?

5-43. *Method of Indicators*

Example : Let $X_1, X_2, \ldots, X_k$ be independent with masses $P\{X_i = j\} = 1/n, j = 1, 2, \ldots, n$, $i = 1, 2, 3, \ldots, k$

If D is the number of **distinct** values among $X_1, X_2, \ldots, X_k$, show that

$$\mathrm{E}(D) = n\left\{1-\left(\frac{n-1}{n}\right)^k\right\} \approx k - \frac{k^2}{2n}, \text{ when } \frac{k^2}{n} \text{ is small.}$$

Solution. Let $p = P\{X_i = j\}$, then $q = P\{X_i \neq j\} = 1 - p$.

$$\therefore \quad P\{X_1 \neq j, X_2 \neq j, \ldots, X_k \neq j\} = q \,.\, q \ldots q = (q)^k \qquad [X_i \text{ are indep}]$$

So $P\{L_1\} = P\{\text{At least one } X_i = j\} = 1 - P\{X_i \neq j \text{ for all } i\} = 1 - (q)k$

Now define the indicator :

$$I_j = \begin{cases} 1, \text{if } X_i = j \text{ for some } i, \text{i.e. for at least one } i \\ 0, \text{otherwise} \end{cases}$$

Then

$$\begin{aligned} \mathrm{E}(I_j) &= 1 \,.\, P(I_j = 1) + 0 \,.\, P(I_j \neq 1), \text{ for some } i \\ &= P(I_j = 1), \text{ for at least one } i \\ &= 1 - (q)^k \end{aligned}$$

By definition of D,

$$D = \sum_n I_j, j = 1, 2, \ldots, n$$

$$\begin{aligned} \therefore \quad \mathrm{E}(D) &= \sum_{j=1}^{n} \mathrm{E}(I_j) = \sum_{j=1}^{n} (1-q^k) = (1-q^k) \sum_{j=1}^{n} 1 = n(1-q^k) \\ &= n\{1-(1-p)^k\} = n\left\{1-\left(1-\frac{1}{n}\right)^k\right\} \\ &= n\left\{1-\left(1-\frac{k}{n}+\frac{k(k-1)}{2}\cdot\frac{1}{n^2}\ldots\right)\right\}, \qquad \text{[By Bin Thm.]} \\ &\approx k - \frac{k^2}{2n} \end{aligned}$$

5-50. *Expectations for Bivariate Distrbutions*

***Def.* 1.** Let X and Y be two variates with a given joint p.d.f. $f(x, y)$ and let g be a real-valued function of two real variables, $g : R^2 \to R$. If the distribution of $Z = g(X, Y)$ is known, then we define

$$\mathsf{E}(Z)=\sum_{z_i} z_i P(Z=z_i),\ Z:\text{discrete.}\quad \mathsf{E}(Z)=\int_{-\infty}^{\infty} z f_Z(z)\,dz,\ Z:\text{continuous}$$

This definition is, in effect, a uni-variate expectation. If f_Z is no longer feasible to investigate, we adopt :

Def. **2.** The expectation of $Z = g(X, Y)$, where X, Y are jointly distributed with p.d.f. $f(x, y)$ is defined by

$$\mathsf{E}(Z)=\mathsf{E}[g(X,Y)]=\begin{cases}\sum_i\sum_j g(x_i,y_j)\,P(X=x_i,Y=y_j) & [X,Y \text{ are discrete}]\\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)\,f(x,y)\,dx\,dy, & [X,Y:\text{continuous}]\end{cases}$$

provided the double sums (or double integrals) are absolutely convergent.

The equivalence of two definitions can be established, as in Art. 5-20 (L.U.S.) and shall not be repeated here.

For the special cases, $g(X, Y) = X$ or Y, $\{P(x_i, y_j) = P(X = x_i, Y = y_j)\mid$ we obviously have

$$\mathsf{E}(X)=\begin{cases}\sum_i\sum_j x_i\,P(X=x_i,Y=y_j)\\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,x\,dx\,dy\end{cases}\qquad \mathsf{E}(Y)=\begin{cases}\sum_i\sum_j y_i\,P(X=x_i,Y=y_j)\\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,y\,dx\,dy\end{cases}$$

5-51. *Two Basic Properties of Expectation*

1. *Expectation of a Linear Combination : Linear Property*

Let g and h be real-valued functions of two real variables x and y. If a and b are any constant, and $Z = ag(X, Y) + bh(X, Y)$ then

$$\mathsf{E}(Z)\equiv\mathsf{E}[ag(X,Y)+bh(X,Y)]=a\mathsf{E}[g(X,Y)]+b\mathsf{E}[h(X,Y)]\qquad[\text{lin } \mathsf{E}]$$

Proof. The result follows instantly, since by the Kernel definition of Expectation

$$\mathsf{E}(Z)\equiv\mathsf{E}[ag(X,Y)+bh(X,Y)]=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}[ag(x,y)+bh(x,y)]\,f(x,y)dy\,dx$$

$$=a\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)f(x,y)\,dy\,dx+b\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x,y)\,f(x,y)\,dy\,dx=a\mathsf{E}[g(X,Y)]+b\mathsf{E}[h(X,Y)].$$

***Cor.* 1.** Let $g(X, Y) = X$, $h(X, Y) \equiv 1$. Then : $\mathsf{E}(aX + b) = a\mathsf{E}(X) + b$.

***Cor.* 2.** Let $g(X, Y) = X$, $h(X, Y) = Y$, then $\mathsf{E}(aX + bY) = a\mathsf{E}(X) + b\mathsf{E}(\mathsf{E}Y)$.

Note that this result is true whether X and Y are independent or dependent. This rule in Cor. 2, being very important, we give an independent proof.

Lin E. Let X and Y be two variates, either both discrete or both absolutely continuous. If $\mathsf{E}(X)$ and $\mathsf{E}(Y)$ exist, then $\mathsf{E}(aX + bY + c)$ exists and

$$\boxed{\mathsf{E}(aX+bY+c)=a\mathsf{E}(X)+b\mathsf{E}(Y)+c}\qquad (a,\ b,\ c \text{ constants})$$

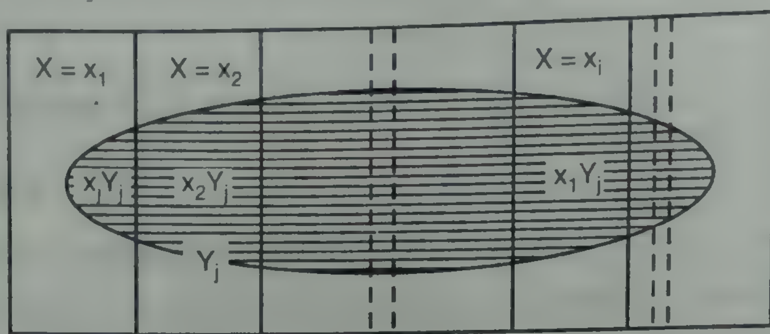
Proof. (I) Assume that both X and Y are discrete. Let $\{x_i\}$ and $\{y_j\}$ be the possible values of X and Y and let us use notation :

$$P(X=x_i,\ Y=y_i)=P(x_i,y_i)=p_{ij};\ \ P(X=x_i)=P(x_i);\ \ P(Y=y_j)=P(y_j)$$

The possible values of $aX + bY + c$ are $ax_i + by_j + c$ with probability, $P(x_i, y_j)$. Observe that,

$$\sum_i P(x_i, y_j) = \sum_i p_{ij} = P\{(x_1, y_j) \uplus (x_2, y_j) \uplus \ldots\} = P(y_j) \quad \ldots(i)$$

Similarly, $\sum_j P(x_i, y_j) = \sum_j p_{ij} = P(x_i)$. ...(ii)



LINEARITY OF EXPECTATION "E"

Since $|ax_i + by_j + c| \le |a||x_i| + |b||y_j||c|$ hence

$$\sum_i \sum_j |ax_i + by_j + c| p_{ij} \le |a| \sum_{ij} |x_i| p_{ij} + |b| \sum_{ij} |y_i| p_{ij} + |c| \Sigma p_{ij}$$

This is equivalent to

$$E(|aX + bY + c)|) \le |a| E(|X|) + |b| E(|Y|) + |c|.1 < \infty$$

because $E(|X|) < \infty$, $E(|Y|) < \infty$. Consequently $E(aX + bY + c)$ exists. Now,

$$\sum_i \sum_j (ax_i + by_i + c) p_{ij} = a \sum_i \sum_j x_i p_{ij} + b \sum_i \sum_j y_j p_{ij} + c \sum_i \sum_j p_{ij}$$

$$= a\Sigma_i x_i [\Sigma_j p_j] + b\Sigma_j y_j [\Sigma_i p_{ij}] + c = a\Sigma_i x_i P(x_i) + b\Sigma_j y_j P(y_j) + c, \quad \text{[by (i) and (ii)]}$$

$$\therefore \quad E[(aX + bY + c)] = aE(X) + bE(Y) + c$$

(II) If X and Y are both absolutely continuous, we see that

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |ax + by + c| f(x, y)\,dxdy \le |a| \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |x| f(x, y)\,dxdy + |b| \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |y| f(x, y)\,dxdy + |c| \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x, y)\,dxdy$$

This is equivalent to : $E|aX + bY + c| \le |a| E(|X|) + |b| E(|Y|) + |c| < \infty$.

because $E(|X|) < \infty$, $E(|Y|) < \infty$. Conseqnetly, $E(aX + bY + c)$ exists. Now,

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} [ax + by + c] f(x, y)\,dxdy = a\int_{-\infty}^{\infty} x\,dx \int_{-\infty}^{\infty} f(x, y)\,dy + b\int_{-\infty}^{\infty} y\,dy \int_{-\infty}^{\infty} f(x, y)\,dx + c\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f\,dx\,dy$$

$$= a\int_{-\infty}^{\infty} x f_X(x)\,dx = b\int_{-\infty}^{\infty} y f_Y(y)\,dy + c$$

$$\therefore \quad E(aX + bY + c) = aE(X) + bE(Y) + c.$$

2. *Expectation of a Product of Independent Variates*

If g and h are real-valued functions of x, y and if X and Y are *independent* variates, then

$$E[g(X)\,h(Y)] = E[g(X)]\,E[h(Y)].$$

Proof. Here $f(x, y) = f_1(x)\,f_2(y)$ [X and Y independent variates] ...(i)

Now $\quad \mathsf{E}[g(X) \cdot h(Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x)\, h(y)\, f(x, y)\, dy\, dx$ [Def. of E]

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x)\, h(y) \,.\, f_1(x)\, f_2(y)\, dy\, dx \qquad \text{[by (i)]}$$

$$= \int_{-\infty}^{\infty} g(x)\, f_1(x)\, dx \,.\, \int_{-\infty}^{\infty} h(y)\, f_{2_1}(y)\, dy = \mathsf{E}[g[(X)] \,.\, \mathsf{E}[h(Y)]].$$

Cor. Let $g(X) = X$ and $h(Y) = Y$, then above result reads : $\mathsf{E}(XY) = \mathsf{E}(X)\,\mathsf{E}(Y)$.

In view of the importance of this Cor., we append an independent proof.

Product Rule. Let X and Y be two variates, either both discrete or both continuous. If $\mathsf{E}(X)$, $\mathsf{E}(Y)$ exist and if X, Y are *independent,* then $\mathsf{E}(XY)$ exists and

$$\mathsf{E}(XY) = \mathsf{E}(X)\,\mathsf{E}(Y)$$

Proof. (I) Assume that both X and Y are discrete. Let $\{x_i\}$ and $\{y_j\}$ be the possible values of X, Y for various i, j. We denote

$P(X = x_i) = P(x_i)$, $P(Y = y_j) = P(y_j)$; $P(X = x_i, Y = y_j) = P(x_i, y_j) = P(x_i) \,.\, P(y_j)$, [by independence]

Now $\mathsf{E}(X) < \infty$, $\mathsf{E}(Y) < \infty$; so, using independence.

$$\sum_i \sum_j | x_i y_j | P(x_i, y_j) = [\sum_i | x_i | P(x_i)]\,[\sum_j | y_j | P(y_j)] = \mathsf{E}\,|X|\,.\,\mathsf{E}\,|Y| < \infty.$$

It follows that $\mathsf{E}(XY) < \infty$, i.e. it exists. Further, using independence.

$$\Sigma_i \Sigma_j x_i y_j\, P(x_i, y_j) = [\Sigma_i x_i P(x_i)]\,[\Sigma_j y_j P(y_j)] \Rightarrow \mathsf{E}(XY) = \mathsf{E}(X) \,.\, \mathsf{E}(Y)$$

(II) When X and Y are both continuous, the proof is similar to (i), save that integrations replace summations.

Extension. If $X_1, X_2, \ldots, X_n$ are mutually *independent* variates, then

$$\mathsf{E}(X_1, X_2, \ldots, X_n) = \mathsf{E}(X_1)\,\mathsf{E}(X_2) \ldots \mathsf{E}(X_n).$$

The proof is immediate by the principle of Mathematical induction.

5-52. *Property.* $\quad \overline{F}_X(t) \geq \overline{F}_Y(t)$ for all $t \in \mathbb{R}$, then $\mathsf{E}(X) \geq \mathsf{E}(Y)$

Proof. $\mathsf{E}(X) - \mathsf{E}(Y) = \mathsf{E}(X - Y) = \int_0^{\infty} [\overline{F}_X(t) - \overline{F}_Y(t)]\, dt + \int_{-\infty}^{0} \{[1 - F_X(t)] - [1 - F_t(y)]\}\, dt$ [§5-11]

$$= \int_{-\infty}^{\infty} [\overline{F}_X(t) - \overline{F}_Y(t)]\, dt \geq 0, \text{ by hypothesis}$$

This gives $\mathsf{E}(X) \geq \mathsf{E}(Y)$.

Comment. We say, X is *randomly larger* than Y if $\overline{F}_X(t) \geq \overline{F}_Y(t), \forall\, t$.

5-60. *Variance, Covariance and Correlation Coefficient*

Let X and Y be any two r.v.s $\mathsf{E}(X) = \mu_X$ and $\mathsf{E}(Y) = \mu_Y$. The following quantities are useful :

(a) The *Variance* of a variate X is denoted by var (X) or σ_X^2 and is defined by

$$\boxed{\mathrm{Var}\,(X) = \mathsf{E}(X - \mu_X)^2 = \mathsf{E}(X^2) - \mathsf{E}^2(X)}$$

(b) The *Covariance* between X and Y is denoted by cov (X, Y) or σ_{XY} and is defined by

$$\boxed{\text{Cov}(X,Y) = E[(X-\mu_X).(Y-\mu_Y)] = E(XY) - E(X)E(Y)}$$

(c) The Correlation Coefficient between X and Y is denoted by Corr (X, Y) or $\rho(X, Y)$ and defined by

$$\boxed{\rho(X,Y) = \sigma_{XY}/\sigma_X\sigma_Y}$$

Sometimes we write $r_{XY} = \rho_{XY}$ or simple r instead of $\rho(X, Y)$.
Some analysis and application of these concepts are presented in section 5-70 onwards.

5-61. *Theorem*

If $E(XY) = E(X)\,E(Y)$, i.e. $\sigma_{XY} = 0$, then X and Y need not be independent random variables.
Proof. We consider a counter example, many more counter examples are possible. Suppose X has a symmetric density f that is, $f(-x) = f(x)$. Let $Y = |X|$. Now

$$E(X) = \int_{-\infty}^{\infty} xf(x)\,dx = 0;\ E(XY) = \int_{-\infty}^{\infty} x.|x|\,f(x)\,dx = 0 \qquad \text{[Odd Integrand]}$$

Thus, $E(XY) = 0 = E(X)\,E(Y)$. But Y is related to X by $Y = |X|$; hence X and Y are not independent.
Note. (i) Let $X \sim$ unif $(-1, 1)$; then $f(x) = 1.2$, $-1 \le x \le 1$. Now let $Y = X^2$.

$$E(X) = \int_{-1}^{1} x.\tfrac{1}{2}\,dx = 0, E(XY) = E(X^3) = \int_{-1}^{1} x^3.\tfrac{1}{2}\,dx = 0.$$

Thus $E(XY) = E(X)\,E(Y)$, although $Y = X^2$, i.e. X and Y are dependent.

(ii) Let $p(-1) = \frac{1}{4}$, $p(0) = \frac{1}{2}$, $p(1) = \frac{1}{4}$; Let $Y = X^2$

$$E(X) = (-\tfrac{1}{4}) + 0 + (\tfrac{1}{4}) = 0 = E(XY) = E(X^3) = 0$$

Thus $E(XY) = E(X)\,E(Y)$, although $Y = X^2$ i.e. X and Y are dependent.

5-62. *Worked-out Problems*

***Example* 1.** Discrete variates X, Y are so distributed that each X and Y have at most two mass points. Prove or disprove : "X and Y are independent iff they are uncorrelated".
Solution. Let X assume only two non-zero values a and b, $(a \ne b)$ and Y assume only two non-zero values c and d, $(c \ne d)$. Write $A = \{X = a\}$, $C = \{Y = c\}$ and suppose $\mathbb{I}_A$ and $\mathbb{I}_C$ are their indicators so that $E(\mathbb{I}_A) = P(A)$ and $E(\mathbb{I}_C) = P(C)$. Then we can express X and Y as under :

$$X = (a-b)I_A + b,\ \ Y = (c-d)I_C + d \qquad [I_A = 1 \text{ if } X = a, \text{etc.}]$$

$$\therefore \qquad \text{Cov}(X, Y) = \text{Cov}\,[(a-b)I_A + b, (c-d)I_C + d]$$

$$= (a-b)(c-d)\,\text{Cov}\,(I_A, I_C)$$

$$= (a-b)(c-d)\,\{E(I_A I_C) - E(I_A)\,E(I_C)\} \qquad [I_A I_C = I_{AC}]$$

$$= (a-b)(c-d)\,\{P(AC) - P(A)\,P(C)\}$$

Thus Cov $(X, Y) = 0$ iff $P(AC) = P(A)\,P(C)$

Note that Ind $(A, C) \Rightarrow$ Ind $(\bar{A}, C)$, Ind $(A, \bar{C})$, Ind $(\bar{A}, \bar{C})$, etc.

***Example* 2.** Let X and Y be two variates having finite means. Prove or disprove :
(i) $E\{\min(X, Y)\} \le \min[E(X), E(Y)]$ (ii) $E[\max(X, Y)] \ge \max\{E(X), E(Y)\}$
(iii) $E[\min(X, Y) + \max(X, Y)] = E(X) + E(Y)$

Solution. We recall the definitions :

$$\min(X, Y) = \tfrac{1}{2}[(X+Y) - |X-Y|]; \ \max(X, Y) = \tfrac{1}{2}[(X+Y) + |X-Y|]; \quad \ldots(1)$$

Also $\quad |E(X-Y)| \le E(|X-Y|)$ [By Ex. 5-2] ...(2)

(i) $E[\min(X, Y)] = \tfrac{1}{2}[E(X+Y) - \tfrac{1}{2}|E(|X-Y|) \le \tfrac{1}{2}E(X+Y)|E(X-Y)|]|$ [by 1(b) and (2)]

i.e. $E(\min(X, Y) \le \tfrac{1}{2}[\{E(X) + E(Y) - |E(X) - E(Y)|\} = \min[E(X), E(Y)]$ [by 1(b)]

This (i) is true.

(ii) $E[\max(X, Y)] = \tfrac{1}{2}[E(X+Y) + E|X-Y|)] \ge \tfrac{1}{2}E(X+Y)|E(X-Y)|]|$. [by 1(b) and (2)]

$= \tfrac{1}{2}[E(Y) + |E(X) - E(Y)|] = \max[E(X), E(Y)]$. [by 1(b)]

Thus $E[\max(X, Y)] \ge \max\{E(X), E(Y)]$ is true.

(iii) Obviously : $\max(X, Y) + \min(X, Y) = X + Y$.

$\therefore \quad E[\max(X, Y) + \min(X, Y)] = E(X+Y) = E(X) + E(Y)$. [True]

***Example* 3.** Prove or disprove :
(a) $E(Y) > E(X)$ iff $F_X(t) > F_Y(t), \forall t$. (b) If $F_X(t) = F_Y(t) \ \forall t$, then $P(X = Y) = 1$, i.e. identically distributed variates are certainly equal. (c) If $E(Y) > E(X)$, then $F_X(t) > F_Y(t)$ for some t.

Solution. $\quad E(Y) = \int_0^{\infty}[1 - F_Y(t)]\,dt - \int_{-\infty}^{0} F_Y(t)]\,dt \quad$ [§ 5-11]

$\therefore \quad E(Y) - E(X) = \int_{-\infty}^{\infty}[F_X(t)] - F_Y(t)]\,dt.$ [On simplification]

Thus, $\quad F_X(t) \ge F_Y(t) \Rightarrow E(Y) > E(X), \forall t$.

Now consider some c.d.f.'s of X and Y defined as under :

$$F_X(t) = \begin{cases} \tfrac{1}{2}, & 0 \le t < 1 \\ 1, & t \ge 1 \end{cases} \quad F_Y(t) = \begin{cases} \tfrac{3}{4}, & 0 \le t < 4 \\ 1, & t \ge 4 \end{cases}$$

Obviously, $F_Y(t) > F_X(t), \forall t$. The corresponding expectations are

$$E(X) = \int_0^{\infty}[1 - F_X(t)]\,dt = \int_0^1 \tfrac{1}{2}\,dt + \int_1^{\infty}(1-1)\,dt = \tfrac{1}{2}$$

$$E(Y) = \int_0^{\infty}[1 - F_Y(t)]\,dt = \int_0^4 \left(1 - \tfrac{3}{4}\right)dt + \int_4^{\infty}(1-1)\,dt = 1.$$

Thus $E_Y(t) > E_Y(t) \Rightarrow F(Y) > F(X)$. It follows that (a) is untrue. Vide Preperty §5-52.

(b) We define X, Y on $\Omega = \{h, t\}$, coin-outcomes, by cross-images :

$X : \Omega \to [0, 1]; P\{X(h) = 0\} = \tfrac{1}{2} = P\{X(t) = 1\}; Y : \Omega \to [0, 1]; P\{Y(h) = 1\} = \tfrac{1}{2} = P\{Y(t) = 0\}$

Obvisouly, $\quad F_X(z) \equiv F_Y(z) = \begin{cases} 0, & z < 0 \\ \tfrac{1}{2}, & 0 \le z < 1 \\ 1, & z \ge 1. \end{cases}$

Thus, X and Y are identically distributed, but $P\{\omega : X(\omega) = Y(\omega)\} = P(\varnothing) = 0$.

Note. (i) $P(X = Y) = P\{(X = 1, Y = 1) \uplus (X = 0, Y = 0\} = 0$ as $(X = 1, Y = 1) = \varnothing = (X = 0, Y = 0)$. (ii) Other examples are also possible. If X is $N(0, 1)$, then X and $-X$ are identically distributed, but obviously, they are unequal.

(c) $\mathrm{E}(Y) - \mathrm{E}(X) = \int_{-\infty}^{\infty} [F_X(t) - F_Y(t)]\, dt = J$ [by (a)]

Thus $\mathrm{E}(Y) > \mathrm{E}(X) \Rightarrow J > 0 \Rightarrow F_X(t) > F_Y(t)$, atleast for some t.

***Example* 4.** Prove or disprove

(a) $P(X > Y) = 1$ iff $\mathrm{E}(X) > \mathrm{E}(Y)$, (b) $\mathrm{E}(X) > \mathrm{E}(Y) \Rightarrow P(X > Y) > \frac{1}{2}$.

Solution. (a) Let $Z = X - Y$; then $\mathrm{E}(Z) = \Sigma z f(z) = \mathrm{E}(X) - \mathrm{E}(Y)$.

By hypothesis, $p(Z > 0) = 1$, it follows that Z assumes only non-negative values; so $\Sigma z f(z) \geq 0$, whence $\mathrm{E}(Z) \geq 0 \Rightarrow \mathrm{E}(X) > \mathrm{E}(Y)$. Thus $P(X > Y) = 1 \Rightarrow \mathrm{E}(X) > \mathrm{E}(Y)$.

Note. $P(X \geq Y) = 1 \Rightarrow \mathrm{E}(X) \geq \mathrm{E}(Y)$ is true.

Now consider the bivariate distribution as shown

$y \rightarrow$ / $x \downarrow$	1	2	3	$P(x)$
2	p	p	0	$2p$
3	p	p	$2p$	$4p$
$P(y)$	$2p$	$2p$	$2p$	$6p$

$$\mathrm{E}(X) = 4p + 12p = 16p = \tfrac{8}{3}$$

$$\mathrm{E}(Y) = 2p + 4p + 6p = 12p = 2.$$

Thus $\mathrm{E}(X) > \mathrm{E}(Y)$. Now

$\{X > Y\} = \{(2, 1), (3, 1), (3, 2)\}$, $P(X > Y) = 3p = \frac{1}{2}$.

Thus $\mathrm{E}(X) > \mathrm{E}(Y) \Rightarrow P\{(X > Y\} = 1$. So (a) is untrue.

(b) Recall : $\mathrm{E}(X) = \int_0^{\infty} [1 - F_X(x)]\, dx - \int_{-\infty}^{0} F_X(x)\, dx$

$$\therefore \quad \mathrm{E}(X) - \mathrm{E}(Y) > 0 \Rightarrow \int_{-\infty}^{\infty} [F_Y(t) - F_X(t)]\, dt > 0 \Rightarrow F_Y(t) > F_X(t) \quad [F \text{ is monotone}] \quad ...(1)$$

$$P(X > Y) = \int_{-\infty}^{\infty} P(X > Y \mid Y = y)\, f_2(y)\, dy = \int P(X > y)\, f_2(y)\, dy$$

$$= \int_{-\infty}^{\infty} [1 - P(X \leq y)]\, f_2(y)\, dy = 1 - \int_{-\infty}^{\infty} F_X(t)\, f_2(t)\, dt > 1 - \int_{-\infty}^{\infty} F_Y(t) f_2(t)\, dt = 1 - \frac{1}{2} = \frac{1}{2}, \quad \text{[by (1)]}$$

In above integral, we have used $F(-\infty) = 0$, $F'(\infty) = 1$, $F(t) = f(t)$. Thus (b) holds.

Problems with Solutions Provided at the End of the Text

1*. Let X and Y be dependent variates. Find the minimum value of
$$H = \mathrm{E}[Y - (a + bX)]^2.$$

2*. Show that if $\mathrm{E}(XY) = \mathrm{E}(X)\,\mathrm{E}(Y)$, then X and Y need not be independent.

3*. n identical coins are tossed. What is the expectation of the number of heads ?

4*. A deck of cards is well-shuffled. What is the expectation of the number of cards on top and the first ace from the top ?

5*. Three urns contain respectively 3 green, 2 white balls; 5 green and 6 white balls; 4 white and 2 green balls. One ball is drawn from each urn. Find the expected number of white balls drawn out.

6*. An urn contains a white and b black balls and c balls are drawn. Find the mathematical expectation of the number of the white balls drawn.

7*. *Die Distribuition.* Two distinct integers are chosen at random and without replacement, from the first six positive integers. Compute the expected value of the absolute difference of these two numbers.

8*. In a certain society, an individual pays income tax only if his income x is greater than a, the amount of tax being $c(x - a)$ where $0 < c < 1$. The distribution of the incomes of individuals liable to tax has density function $f(x) = k/(x)^{\theta + 1}$, $x > a$ $\theta > 1$; $f(x) = 0$, otherwise. Show that the average tax paid is $ca(\theta - 1)$.

It is proposed to increase the tax on each individual whose income x is greater than $b(b > a)$ by an amount $c(x - b)$. Show that this will increase the average tax paid by 20%, provided b is chosen so that $(b/a)^{\theta - 1} = 5$.

9*. Find $\mathrm{E}(XY)$ and $\mathrm{E}(Y^2/X)$, when variates X and Y have the joint p.d.f.
$P(X = x, Y = y) = (x + y^2)/42$, $x = 1, 4$; $y = -1, 0, 1, 3,$; $P(x, y) = 0$, elsewhere.

10*. For any constants a and b, show that the variates X and Y are independent iff

$$\mathrm{E}[u(a - X)\, u(b - Y)] = \mathrm{E}[u(X - a)] \,.\, \mathrm{E}[u(b - Y)] \qquad \ldots(1)$$

where $u(z)$ is the *unit step function* defined by $u(z) = 1$, if $z \geq 0$; $u(z) = 0$, if $z < 0$.

11*. Events $A_1, A_2, \ldots, A_n$ are mutually independent iff their respective indicator variates $J_1, J_2, \ldots, J_n$ are independent.

Exercise 5(b)

1. (a) A die is thrown. Show that the expectation of the number on this die is 7/2.

   (b) Two unbiased dice are thrown. Show that the expected values of the sum and product of numbers of points on them are 7 and 49/4.

   (c) n unbiased dice are thrown. Show that the expected values of the sum and product of numbers of points on these dice are $7n/2$ and $(7/2)^n$.

2. Two cards are drawn at random from 10 cards numbered 1 to 10. Show that the expectation of the sum of points on two cards is 11.

3. In a lottery m tickets are drawn out of n tickets numbered from 1 to n. Show that the expectaiton of the sum of cubes of the numbers on the tickets drawn is $mn(n + 1)^2/4$.

4. X and Y have the joint discrete distributions

(i)

$X\downarrow$ \ $Y\rightarrow$	-1	2	4
-6	$2p$	$2p$	p
6	$4p$	p	$2p$

(ii)

$X\downarrow$ \ $Y\rightarrow$	0	2	3
-1	p	$2p$	p
2	p	p	$2p$
0	p	p	0

(i) Find p. If $Z = 3 + 6X + 2Y$. Show that $\mathrm{E}(Z) = 11 = \mathrm{E}(Z^2) = 1471$.

(ii) Find p. Show that $\mathrm{E}(X^2Y) = 39/10$ and $\mathrm{E}[X/(Y + 1)] = 7/40$.

5. A bag contains a coin of value M and a number of other coins whose aggregate value is m. A person draws coins one at a time till he gets the coin M. Show that the value of his expectation is $M + (m/2)$.
6. Balls are tkaen one by one out of an urn containing w white balls and b black balls until the first white ball is drawn. Prove that the expectation of the number of black balls preceding the first white ball is $b/(w + 1)$.
7. A penny and dime are tossed and suppose X denotes the number of heads turned up. Then the penny is tossed again. Let Y denote the number of heads up on the dime (from the first toss) and the penny from the second toss.
   (a) Find the joint and marginal p.d.f. of X and Y.
   (b) Find the conditional distribution of Y given $X = 0, 1, 2$.
   (c) Show that Corr $(X, Y) = \frac{1}{2}$. **[Ans.** $f(0) = f(2) = \frac{1}{4}, f(1) = \frac{1}{2}; g(0) = g(2) = \frac{1}{4}, g(1) = \frac{1}{2}$]
8. You order a wall to wall carpet $m \times n$ foot rectangular dim. The rug should be cut in one piece from a very wide (more than m feet) and long roll of material. It is known that the workman cutting the carpet never gives less than the ordered amount but usually makes a 'random error' in the measurement of each linear dimension, which is, on average 2% of the required length. Assuming the errors in measuring the length and width of a carpet to be independent, what is the expected area of the delivered carpet ?
9. An urn contains three balls numbered 1, 2, 3 and a sample of size two is drawn without replacement from it. Let X be the number on the first ball draw, Y the larger of the two numbers drawn and Z the smaller of the two numbers drawn.
   (a) Find the joint discrete density of (i) X and Y, (ii) Y and Z.
   (b) Show that $P(X = 1 \mid Y = 3) = 1/4$, Corr $(X, Y) = \sqrt{3}/4$, Corr $(Y, Z) = 1/2$.
   (c) Show that the conditional distribution of Y is $P(Y = 2 \mid Z = 1) = \frac{1}{2} = P(Y = 3 \mid Z = 1)$
10. An urn contains 4 balls, two of the balls are numbered with a 1 and the other two are numbered with a 2. Two balls are drawn from the urn without replacement. Let X and Y denote the smaller and larger of the number on the drawn balls. Find the joint p.d.f. of X and Y and show that the marginal densities of X and Y are $f(1) = g(2) = 5/6, f(2) = g(1) = 1/6$. Also Corr $(X, Y) = 1/5$.
11. The joint p.d.f. is $f(x, y) = 24y(1 - x), 0 \le y \le x \le 1$; $f(x, y) = 0$, elsewhere.
    Show that $E(X^2Y) = 4/21$, $E(Y^2X) = 1/7$ and $E[Y^2/(1 - X)] = 6/5$.
12. The joint p.d.f. of X and Y is given by $f(x, y) = 2x(x - y), -x < y < x$.
    Find the marginals and show that $E(X^nY^m) = \dfrac{2}{m+n+4}\left\{\dfrac{1+(-1)^m}{m+1} - \dfrac{1+(-1)^{m+1}}{m+2}\right\}$.
13. X and Y are jointly continuous and $f(y \mid x) = 1, x < y < x + 1; f_X(x) = 1, 0 < x < 1$. Find $P(X + Y < 1)$, $f(x \mid y)$ and show that $E(Y) = 1$, and Corr $(X, Y) = 1/\sqrt{2}$.
14. The joint p.d.f. of X and Y is $f(x, y) = x + y, 0 \le x, y \le 1$; $f(x, y) = 0$, otherwise. Show that $E\{\min(X, Y)\} = 5/12$ and $E\{\max(X, Y)\} = 9/12$.
15. Let $X_1, X_2, ..., X_n$ be a sequence of *i.i.d.* variables. Suppose X_k assumes only positve integral values and $E(X_k) = \mu$ exists, $1 \le k \le n$. Let $S_n = X_1 + X_2 + ... + X_n$.
    (a) Show that $E(S_m/S_n) = m/n$, for $1 \le m \le n$.
    (b) Show that $E(S_n^{-1})$ exists, and $E(S_m/S_n) = 1 + (m - n)\mu E(S_n^{-1})$ for $1 \le m \le n$.

(c) Verify and use the inequality $x + x^{-1} \geq 2, (x > 0)$ to show $E(S_m/S_n) \geq (m/n)$, for $m, n \geq 1$.

16. Show that the expected length of the chords of a circle of radius a taken at random in the circle is $4a/\pi$.

17. Two points P and Q are selected at random in a square of side (length) a. Show that $E(|PQ|^2) = a^2/3$.

18. Two points $P(r_1, \theta_1)$ and $Q(r_2, \theta_2)$ are selected uniformly at random in a circle of radius a and centre at origin. Show by brutal integration that $E(|PQ|) = 128a/45\pi$. Verify by Crofton method.

19. The probability of a shot hitting a point distant r from the centre O of the target follows the p.d.f. $p(r) = 2/\pi(1 + r^2)$. The target is divided into four regions as follows :

C_1 : within the circle with centre O and radius $1/\sqrt{3}$

C_2 : without C_1 but within the concentric circle of radius 1.

C_3 : without C_2 but within the concentric circle of radius $\sqrt{3}$; C_4 : without C_3.

The scores for hits in the regions C_1, C_2, C_3, C_4 are 4, 3, 2, 0 respectively. Let X_i be the number of hits registered in the region C_i, $i = 1, 2, 3, 4$ when a total of N shots is fired. Find :

(a) The probability distribution of X_1, X_2, X_3, X_4.

(b) The probability that the score after 5 shots exceeds 15.

(c) Expected score after 5 shots.

20. If $X_1, X_2, ..., X_n$ is a set of positive interchangeable variates, show that $E(S_k / S_n) = k/n,\ 1 \leq k \leq n;\ S_n = X_1 + X_2 + ... + X_n$.

[Def. Set of variates $X_1, X_2, ..., X_n$ is said to be *interchaneable* if the joint p.d.f. of $(X_1, ..., X_n)$ remains unaltered under any permutation of the X_i's.
Note that i.i.d. variates are always interchangeable].

5-70. *Variance : Spread of Distribution*

Let X be a random variable. Then, putting $E(X) = \mu$, the variance of X, written Var (X) or σ_X^2, is defined by $\text{Var}(X) = E(X - \mu)^2$. Thus

$$\text{Var } X = \Sigma_x (x-\mu)^2\, p_X(x),\ X \text{ is discrete};\ \ \text{Var }(X) = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\,dx,\ \ X \text{ is continuous.}$$

The variance describes the spread or dispersion of the distribution about the central point μ. It is a useful quantity in the sense that it summarizes information about the shape of a distribution. As $\text{Var}(X) = \Sigma(X - \mu)^2 \geq 0$ (necessarily non-negative), its *positive* square root is called *standard deviation* (S.D.), and is denoted by σ_X.

Note. Some authors write $V(X)$ for Var(X) : a bad practice.

5-71. *Some Operational Properties*

1. $\text{Var}(X) = E(X^2) - [E(X)]^2$ (Working Formula)
2. $\text{Var }(aX + b) = a^2 \text{ Var }(X)$
3. $\text{Var }(aX + bY) = a^2\sigma_X^2 + b^2\sigma_Y^2 + 2ab\sigma_{XY}$

Proof. These properties are proved using definitions and lin E. Write $X_0 = X - \mu_X$, $Y_0 = Y - \mu_Y$.

(1) $$\text{Var}(X) = \mathsf{E}(X-\mu)^2 = \mathsf{E}(X^2 - 2\mu X + \mu^2) = \mathsf{E}(X^2) - 2\mu\mathsf{E}(X) + \mu^2$$
$$= \mathsf{E}(X^2) - 2\mu^2 + \mu^2 = \mathsf{E}(X^2) - [\mathsf{E}(X)]^2 \qquad [\mu = \mathsf{E}(X)]$$

(2) $$\text{Var}\,(aX+b) = \mathsf{E}\{[aX+b) - \mathsf{E}(aX+b)]^2\} = \mathsf{E}\{[(aX+b) - (a\mu+b)]^2\}$$
$$= \mathsf{E}[a^2(X-\mu)^2] = a^2\mathsf{E}(X-\mu)^2 = a^2\,\text{Var}\,(X) \qquad \text{[Lin E and Def. of Var]}$$

Note. $\text{Var}\,(aX) = a^2\,\text{Var}\,(X);\ \text{Var}\,(b) = \text{Var (Constant)} = 0.$

(3) $$\text{Var}\,(aX+bY) = \mathsf{E}\{[(aX+bY) - \mathsf{E}(aX+bY)]^2\} = \mathsf{E}\{[(aX+bY) - a\mu_X + b\mu_Y)]\}^2$$
$$= \mathsf{E}\{[a(X-\mu_X) + b(Y-\mu_Y)]^2\}\{a^2X_0^2 + b^2Y_0^2 + 2ab(X_0, Y_0)\},$$
$$= a^2\mathsf{E}(X_0^2) + b^2\mathsf{E}(Y_0^2) + 2ab\mathsf{E}(X_0Y_0) = a^2\sigma_X^2 + b^2\sigma_Y^2 + 2ab\sigma_{XY} \ [\text{Lin E \& Defs. of } \sigma_X^2 \text{ and } \sigma_{XY}]$$

Note. Var $(aX + bX + c)$ = Var $(AX + bY)$.

Variance of an Indicator Variate

Here $X = \begin{cases} 1, \text{if event } A \text{ occurs with probability } p \\ 0, \text{if event } A \text{ does not occur with probability } q = 1-p \end{cases}$

$\mathsf{E}(X) = 1\,.\,p + 0\,.\,q = p, \mathsf{E}(X^2) = 1^2\,.\,p + 0^2\,.\,q = p.$ Hence

$$\text{Var}(X) = \mathsf{E}(X^2) - \mathsf{E}^2(X) = p - p^2 = p(1-p) = pq.$$

Minimal Property of Variance

$$\mathsf{E}(X-c)^2 = \mathsf{E}\{[(X-\mu) + (\mu-c)]^2\} = \mathsf{E}\{(X-\mu)^2 + 2(\mu-c)(X-\mu) + (\mu-c)^2\}$$
$$= \mathsf{E}(X-\mu)^2 + 2(\mu-c)\mathsf{E}(X-\mu) + (\mu-c)^2, \qquad \text{[by Lin E]}$$
$$= \sigma_X^2 + 2(\mu-c)(\mu-\mu) + (\mu-c)^2 = \sigma_X^2 + (\mu-c)^2, \qquad [\mathsf{E}(X-\mu) = \mu - \mu = 0]$$

Since $(\mu - c)^2 \geq 0$, hence $\mathsf{E}(X-c)^2 - \sigma_X^2 \geq 0 \Rightarrow \sigma_X^2 \leq \mathsf{E}(X-c)^2$.

Thus, Var(X) is the smallest 2nd order moment, i.e. $\mathsf{E}(X - c)^2$ is the minimum when $c = \mu$.

Aliter. Let $H(c) = \mathsf{E}(X-c)^2 = \mathsf{E}(X^2 - 2cX + c^2) = c^2 - 2c\mathsf{E}(X) + \mathsf{E}(X^2)$

$H'(c) = 2c - 2\mathsf{E}(X), H''(c) = 2 > 0\,.\,H'(c) = 0 \Rightarrow c = \mathsf{E}(X) = \mu$. It follows that $H = \mathsf{E}(X - c)^2$ is minimum when $c = \mu$.

Example : Show that if $P(a \leq X \leq b) = 1$, then $\mathsf{E}(X)$ and Var (X) exist and that $a \leq \mathsf{E}(X) \leq b$ and Var $(X) \leq (b-a)^2/4$.

Solution. Since $a \leq X \leq b$, almost certainly, using $f_X(x)$ as p.d.f. of X, we get

$$a\int_a^b f(x)\,dx \leq \int_a^b xf(x)\,dx \leq b\int_a^b f(x)\,dx \Rightarrow a \leq \mathsf{E}(X) \leq b.$$

Thus, $\mathsf{E}(X) = \mu$ exists. Further

$$a \leq X \leq b \Rightarrow a - \tfrac{1}{2}(a+b) \leq X - \tfrac{1}{2}(a+b) \leq b - \tfrac{1}{2}(a+b) \Rightarrow |X - \tfrac{1}{2}(a+b)| \leq \tfrac{1}{2}(b-a)$$

Thus, $E[X - \frac{1}{2}(a+b)]^2 \le \frac{1}{4}(b-a)^2$. By minimal property, Var $(X) < E(X - c)^2$, it follows that $[c = \frac{1}{2}(a+b)]$, Var (X) exists and that $\sigma_X^2 \le E(X-c)^2 \le \frac{1}{4}(b-a)^2$.

5-72. *Theorem. Variance of Degenerate r.v.*

(i) Let X be a discrete variate. Then Var$(X) = 0$, iff X is Degenerate.

(ii) If X be a continuous variate, then Var$(X) \ne 0$.

Proof. (i) $\text{Var}(X) = \Sigma_i (x_i - \mu)^2 p_i = 0$, which can happen iff $x_i = \mu$ for every i. Consequently, Var $(X) = 0$ iff the only possible value of X is μ, in other words, iff $P(X = \mu) = 1$. Observe that

$$\text{Var}(X) = E(X^2) - [E(X)]^2 = c^2 - c^2 = 0 \qquad (\text{if } X = \text{constant} = c)$$

(ii) When x is continuous with $E(X) = \mu$, then

$$\text{Var}(X) = \int_{-\infty}^{\infty} f(x)(x-\mu)^2\,dx = 0, \text{ iff } f(x) = 0 \text{ whenever } x \ne \mu.$$

However, this is not possible for any density function f defined on $]-\infty, \infty[$.

Remark. Var $(X) = 0 \Leftrightarrow P(X = \mu) = 1$, also follows from Chebyshev's inequality. See Chapter 10.

5-73. *Variance of a Linear Combination of Independent Variates*

If $X_1, X_2, \ldots, X_n$ are independent variates, then

$$\text{Var}(a_1X_1 + a_2X_2 + \ldots + a_nX_n) = a_1^2 \text{ Var}(X_1) + \ldots + a_n^2 \text{Var}(X_n) \qquad \ldots(1)$$

Proof. Let $Z = a_1X_1 + \ldots + a_nX_n$; then $E(Z) = a_1E(X_1) + \ldots + a_nE(X_n)$ [Lin E]

$\therefore \quad Z - E(Z) = a_1[X_1 - E(X_1)] + \ldots + a_n[X_n - E(X_n)] = a_1Y_1 + \ldots + a_nY_n$ where $Y_i = X_i - E(X_i)$ $1 \le i \le n$ are centred variates. Now

$$\text{Var}(Z) = E[Z - E(Z)]^2 = E\{(\Sigma a_iY_i)^2\} = E\{(\Sigma a_iY_i)(\Sigma a_jY_j)\} = E\{\Sigma\Sigma a_ia_jY_iY_j\} = \Sigma\Sigma a_ia_jE(Y_iY_j) \ldots(2)$$

Now X_i and X_j are independent, hence so must be Y_i and Y_j whence

$$E(Y_iY_j) = E(Y_i)E(Y_j) = [E(X_i) - E(X_i)].[E(X_j) - E(X_j)] = 0, i \ne j.$$

Using linearity of E in (2), setting $E(Y_iY_j) = 0 \;\; \forall\, ij,\, i \ne j$, and $E(Y_i^2) = \sigma_i^2$ we get

$$\text{Var}(Z) = \Sigma\Sigma a_ia_j\,E(Y_iY_j) = \Sigma a_i^2\sigma_i^2.$$

Cor. $\text{Var}(\overline{X}) = \sigma^2/n$, if X_i are i.i.d. with Var $(X_i) = \sigma^2$.

Proof. Putting $\sigma_1^2 = \ldots = \sigma_n^2 = \sigma^2; a_1 = \ldots = a_n = 1/n$ into (1) we get,

$$\text{Var}(\overline{X}) = (\sigma^2 + \sigma^2 + \ldots + \sigma^2)/n^2 = \sigma^2/n. \quad [\overline{X} = (X_1 + \ldots + X_n)/n]$$

5-74. Variance Conditioned by an Event

The variance of a random variable X conditioned by the occurrecen of the event A is denoted by $\mathrm{Var}(X \mid A)$, and is defined by

$$\mathrm{Var}(X \mid A) = \mathrm{E}\{(X-\mu)^2 \mid A\}, \qquad [\mu = \mathrm{E}(X)]$$

This can be trivially simplified to

$$\mathrm{Var}\,(X \mid A) = \mathrm{E}(X^2 \mid A) - \mathrm{E}^2(X \mid A).$$

Theorem. Let $\{A_i\}$ be a measurable partition of the sample space Ω and $P(A_i) > 0, \forall i$. If $\mathrm{Var}(X)$ exists (i.e. $\sigma_X^2 < \infty$) then, with $\mu = \mathrm{E}(X)$, we have

$$\mathrm{Var}(X) = \Sigma_i\, P(A_i)\,\mathrm{Var}\,(X \mid A_i) + \Sigma_i P(A_i)\ \ [\mathrm{E}(X \mid A_i - \mu]^2$$

Proof. Recall that

$$\mathrm{Var}(X) = \mathrm{E}(X-\mu)^2 = \Sigma_i\, P(A_i)\,\mathrm{E}[(X-\mu)^2 \mid A_i], \qquad \text{[Art. 5-41]}$$

$$= \sum_i \int_{A_i} (x-\mu)^2\, dP = \sum_i \int_{A_i} \{[X - \mathrm{E}(X \mid A_i)] + [\mathrm{E}(X \mid A_i) - \mu]\}^2\, dP \qquad \text{[adjustment]}$$

$$= \sum_i \left\{\int_{A_i} [X - \mathrm{E}(X \mid A_i)]^2 dP + \int_{A_i} [\mathrm{E}(X \mid A_i) - \mu]^2 dP\right\} \qquad \left\{\because \int_{A_i} [X - \mathrm{E}(X_i \mid A_i)] dP = 0\right\}$$

$$= \sum_i \{P(A_i)\,\mathrm{Var}\,(X \mid A_i) + P(A_i)\,[\mathrm{E}(X \mid A_i) - \mu]^2\}.$$

5-75. Worked-out Problems

***Example* 1.** Let X and Y each take on either the value 1 or -1. Let $P\{X = i, Y = j\} = p(i, j)$, $i = \pm 1$, $j = \pm 1$. Suppose that $\mathrm{E}(X) = \mathrm{E}(Y) = 0$. Show that $p(1, 1) = p(-1, -1)$ and $p(1, -1) = p(-1, 1)$.

Let $P = 2p(1, 1)$. Find $\mathrm{Var}(X)$, $\mathrm{Var}(Y)$, $\mathrm{Cov}\,(X, Y)$.

$X\downarrow$ \ $Y\rightarrow$	1	-1	$f_1(x)$
1	a	b	$a+b$
-1	c	d	$c+d$
$f_2(y)$	$a+c$	$b+d$	1

Bivariate Distribuition

Solution. It is convenient to write probability as a, b, c, d (shown herein). Now

$$\mathrm{E}(X) = 0 \Rightarrow (a+b) - (c+d) = 0$$

$$\mathrm{E}(Y) = 0 \Rightarrow (a+c) - (b+d) = 0.$$

Since $a + b + c + d = 1$, these give

$$a + b = c + d = 1/2 \qquad \text{...(i)}$$

$$a + c = b + d = 1/2 \qquad \text{...(ii)}$$

whence $b = c$, i.e. $p(-1, 1) = p(1, -1)$.

From $a + b + c + d = 1$, $a + b + a + c = 1$, we get $a = d$, i.e. $p\,(1, 1) = p\,(-1, -1)$

Second Part. Given $p = 2a$. Now

$\mathrm{Var}\,(X) = \mathrm{E}(X^2) - \mathrm{E}^2(X) = \mathrm{E}(X^2) - 0 = (1)^2\,(a+b) + (-1)^2\,(c+d) = a+b+c+d = 1.$

$\mathrm{Var}\,(Y) = \mathrm{E}(Y^2) - \mathrm{E}^2(Y) = \mathrm{E}(Y^2) - 0 = (1)^2\,(a+c) + (-1)^2\,(b+d) = a+b+c+d = 1.$

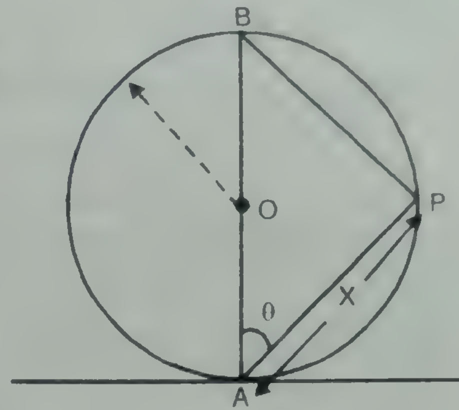
$\mathrm{Cov}\,(X, Y) = \mathrm{E}(XY) - \mathrm{E}(X)\,\mathrm{E}(Y) = \mathrm{E}(XY) - 0$

$= (1)\,(1)\,a + (1)\,(-1)\,b + (-1)\,(1)\,c + (-1)\,(-1)\,d = (a+d) = (b+c)$

$= (a+d) - [1 - (a+d)] = 2(a+d) - 1 = 4a - 1.$

$= 2p - 1.$

***Example* 2.** From a point A on the circumference of a circle of radius a, a chord AP is drawn in a *random* direction. Show that $E(AP) = 40/\pi$ and $\text{Var}(AP) = 2a^2 [1 - (8/\pi^2)]$. Also show that the chance is 1/3 that the length AP will exceed the length of the side of an equilateral triangle inscribed in the circle.



Solution. Let AP be inclined at angle θ with the diameter AOB, then for the random direction of AP, $-\pi/2 \le \theta \le \pi/2$, and θ is uniformly distributed so that $f(\theta) = 1/\pi$.
If $AP = X$, then $X = 2a \cos \theta$. So

$$E(X) = \int_{-\pi/2}^{\pi/2} (2a \cos \theta) \frac{d\theta}{\pi} = \frac{4a}{\pi} \int_0^{\pi/2} \cos \theta \, d\theta = \frac{4a}{\pi}$$

$$E(X^2) = \int_{-\pi/2}^{\pi/2} (4a^2 \cos^2 \theta) \frac{d\theta}{\pi} = \frac{8a^2}{\pi} \int_0^{\pi/2} \cos^2 \theta \, d\theta = 2a^2.$$

$$\text{Var}(X) = \text{Var}(X^2) = E(X^2) - E^2(X) = 2a^2 - (4a/\pi)^2 = 2a[1 - (8/\pi^2)].$$

Now length of the side of the equilateral triangle inscribed in a circle of radius a is $a\sqrt{3}$ (Draw Fig. to check it). So

$$P\{X > a\sqrt{3}\} = P\{\cos \theta > \sqrt{3}/2\} = P\{-\pi/6 \le \theta \le \pi/6\} = \int_{-\pi/6}^{\pi/6} \frac{d\theta}{\pi} = \frac{2}{\pi} \int_0^{\pi/6} d\theta = \frac{1}{3}.$$

***Example* 3.** If $X_1, X_2, \ldots, X_n$ are i.i.d. unif $(0, \beta)$ variates, find the var Y_n.

Solution. The p.d.f. of $Y_n = \max \{X_1, \ldots, X_n\}$, is given by

$$g_n(y_n) = nf(y_n)\left[\int_{-\infty}^{y_n} f(x)\, dx\right]^{n-1} \qquad \left[f(x) = \frac{1}{\beta}, 0 < x < \beta\right]$$

$$= n \cdot \frac{1}{\beta}\left[\int_0^{y_n} \frac{1}{\beta} ax \, dx\right]^{n-1} = \frac{n}{\beta^n} (y_n)^{n-1}$$

For neatness, write y for y_n, then

$$E(Y_n) = \frac{n}{\beta^n} \int_0^{\beta} y^n \, dy = \frac{n}{n+1} \beta, \quad E(Y_n^2) = \frac{n}{\beta^n} \int_0^{\beta} y^{n-1} y^2 \, day = \frac{n}{n+1} \beta^2$$

$$\text{Var}(Y_n) = \frac{n}{n+2} \beta^2 - \left(\frac{n}{n+1}\right)^2 \beta^2 = \frac{n\beta^2}{(n+1)^2 (n+2)}$$

Note. (i) Since $E\{(n+1)Y_n/n\} = \beta$, we state that $(n+1)Y_n/n$ is an unbiased estimation of parameter B.

(ii) $\text{Var}\left(\frac{n+1}{n} Y_n\right) = \frac{(n+1)^2}{n^2} \text{Var}(Y_n) = \frac{(n+1)^2}{n^2} \frac{n\beta^2}{(n+1)^2 (n+2)} = \frac{\beta^2}{n(n+2)}.$

***Example* 4.** If $X \sim$ unif $(a-\frac{1}{2}b, a+\frac{1}{2}b)$, find Var (Y_k).

Solution. $f(x)=\dfrac{1}{b}$, $F(x)=\int_{x_0}^{x}\dfrac{1}{b}dt=\dfrac{x-(a-b/2)}{b}$ $\quad [x_0=a-(b/2),\ x_1=a+(b/2)]$

$$f_k(x)=\frac{n!}{(k-1)!\,(n-k)!}[F(x)]^{k-1}.f(x).[1-F(x)]^{n-k}$$

$$=\frac{n!}{(b)^n\,(k-1)!\,(n-k)!}.(x-x_0)^{k-1}(x_1-x)^{n-k}$$

$$E[X_{(k)}^r]=\frac{n!}{(b)^n\,(k-1)!\,(n-k)!}\int_{x_0}^{x_1}x^r(x-x_0)^{k-1}(x_1-x)^{n-k}dx$$

$$=\frac{n!}{(k-1)!\,(n-k)!}\int_0^1(x_0+bt)^r t^{k-1}(1-t)^{n-k}dt \qquad [x-x_0=bt]$$

$$E[X_{(k)}]=\frac{n!}{(k-1)!\,(n-k)!}\{x_0B(k,n-k+1)+bB(k+1,n-k+1)\} \qquad (r=1)$$

$$=\frac{n!}{(k-1)!\,(n-k)!}\left\{\left(a-\frac{b}{2}\right)\frac{(k-1)!\,(n-k)!}{n!}+b.\frac{k!\,(n-k)!}{(n+1)!}\right\}$$

$$=[a-(b/2)]+[bk/(n+1)]$$

$$E(X_{(k)}^2]=\frac{n!}{(k-1)!\,(n-k)!}\int_0^1(x_0^2+2bx_0t+b^2t^2)\,t^{k-1}(1-t)^{n-k}dt \qquad (r=2)$$

$$=\frac{n!}{(k-1)!\,(n-k)!}\left\{x_0^2B(k,n-k+1)+2bx_0B(k+1,n-k+1)+b^2B(k+2,n-k+1)\right\}$$

$$=x_0^2+2bx_0k/(n+1)+b^2k(k+1)/(n+1)(n+2)$$

$$\text{Var }(X_k)=E(X_{(k)}^2)-[E(X_{(k)})]^2=k(n-k+1)b^2/(n+2)(n+1)^2.$$

Problems with Solutions Provided at the End of the Text

1*. If the c.d.f. of a variate X is $F_X(x)=(1-pe^{-\lambda x})\,I_{[0,\infty)}(x)$, compute Var X.

2*. The distribution of X is given by $P(X=0)=P(X=2)=p$, $P(X=1)=1-2p$, for $0\le p\le\frac{1}{2}$. For what value of p, is the Var (X) maximum ?

3*. *Random Walk Problem.* Starting from the origin, unit steps are taken to the right with probability p and to the left with probability $q(=1-p)$. Assuming independent movements, find the mean and variance of the distribution moved from origin after n steps.

4*. Let X be a random variable which assumes values 1, 2, 3, ... with probability $P(X=k)=pq^{k-1}$, $k=1, 2, \ldots$. Find E(X) and Var(X).

5*. One die is thrown until a four is obtained. Find the most probable number of throws and also the expectation and variance of throws.

6*. A point P is taken at random in a line AB of length $2a$, all positions of the point being equally likely. Show that the expected value of the area of the rectangle $AP.\ PB$ is $2a^2/3$ and that the probability of the area exceeding $a^2/2$ is $1/\sqrt{2}$. Find Var(S) also.

7*. The two equal sides of an isosceles triangle are of length a each, and the angle θ between them has a p.d.f. proportional to $\theta(\pi - \theta)$, in the range $(0, \frac{1}{2}\pi)$ and zero otherwise. Find the mean value and variance of the area of the triangle.

8*. Show that the standard error of the number of successes in a large random sample, when probability of success differs at each draw is less than the standard error of number of successes of the sample (of same size) when probability of success remains constant.

Exercise 5(c)

1. (a) Show that the mean and variance of the following p.m.f. are 3/2 and 3/4 respectively: $f(x) = 2/3^x, x = 1, 2, 3, \ldots$

(b) Let a r.v. X assume n values $x_1, x_2, \ldots, x_n$ such that $a \le x_j \le b$ for all j. Show that $\text{Var}(X) \le (b-a)^2/4$.

(c) Let X represent the difference between the number of heads and the number of tails when a fair coin is tossed 6 times. Find p.m.f. of X and show that Var $(|X|) = 159/64$.

2. If $f(x) = 1/\pi, 0 \le x \le \pi$, show that

(i) $E(\sin X) = 2/\pi$, (ii) $\text{Var}(\sin X) = (\pi^2 - 8)/2\pi^2$.

3. X and Y are independent variates with p.d.f. $P(X = \theta \pm 1) = 1/2$; $P(Y = \theta \pm 2) = 1/2$

Let $Z = aX + bY$ with $E(Z) = \theta$. Find the values of a and b which minimize Var (Z) and show that min Var $(Z) = 4/5$.

4. If t is any positive real number show that the function defined by

$p(x) = e^{-t}(1-e^{-t})^{x-1}, x = 1, 2, 3, \ldots$.

can represent a p.d.f. of a variate X. Show that the mean and variance of the distribution are e^t and $e^t(e^{t-1})$.

5. Show that the mean and variance of the Weibull distribution $f(t) = 2at\, e^{-at^2}, 0 \le t < \infty$ are

$\frac{1}{2}(\pi/a)^{1/2}$ and $(1 - \frac{1}{4}\pi)a^{-1}$ respectively.

6. A continuous variate X has the p.d.f. $f(x) = a + bx, 0 \le x \le 1$. If the mean of the distribution is 1/2, find a and b and show that Var $(X) = 1/12$.

7. Determine k so that the following curve reprsents a frequency function

$f(x) = k(x-1)(2-x)(3-x), 1 < x < 2; f(x) = 0$, elsewhere.

Find the mean and variance of X. What is the probability that a value of X taken at random will exceed 1.8 ? [**Ans.** $k = 4, \mu = 22/15, \sigma^2 = 11/225, p = 0.076$]

8. The c.d.f. of a variate X is given by

$F(x) = 0, x < 0; F(x) = x/8, 0 \le x < 2; F(x = (1/16)x^2, 2 \le x < 4; F(x) = 1, x \ge 4$. Show that $E(X) = 31/12$ and Var $(X) = 167/144$.

9. A random variable X has the p.d.f.
$f(x) = 30x^4(1-x),\ 0 \le x \le 1;\ f(x) = 0$, otherwise.
Show that $P(X \ge 3/4) = 1 - 4(3/4)^7$ and $\text{Var}(X) = 5/196$.

10. A continuous variate has the p.d.f. (Indicator form)
$f(x) = (x/2)\, I(0 \le x < 1) + (1/2)\, I(1 \le x \le 2) + [(3-x)/2]\, I(2 \le x \le 3)$.
Verify that the area under the curve is unity and sketch the graph. Show that the mean is 3/2 and variance of X is 5/12.

11. A continuous variate has the p.d.f. $f(x) = k(a^2 + x^2), -a \le x \le a\sqrt{3}$.
where k is a constant. Find k, μ and σ when $a = 7\pi$. Also obtain the distribution function and evaluate $P(X \ge 0)$. [**Ans.** $k = 12, \mu = 6 \log 2, \Sigma(X^2) = a(12 + 12\sqrt{3} - (7\pi/12), p = 4/7]$

12. Two cards are selected at random with replacement from a box which contains four cards numbered 1, 1, 2 and 2. Let X denote the sum of the numbers shown on the two cards. Find the distribution of X. Also find $E(X)$ and $\text{Var}(X)$.

13. For a variate X, $E(X) = 10$ and $\text{Var}(X) = 25$. Find the positive values of a and b such that $Y = aX - b$ has expectation zero and variance 1. [$a = 1/5$ and $b = 2$].

14. Show that the mean and variance of the number of successes in a series of n independent trials, the probability of success in the ith trial being p_i $1 \le i \le n$ are Σp_i and $\Sigma p_i q_i$ respectively.

15. In an objective type examination, consisting of 50 questions, for each question there are four answers of which only one is correct. A candidata scores 1, if he picks up the correct answer, and $-1/3$, otherwise. If a candidata makes only a random choice in respect of the 50 questions, find his expected score and the variance of his score. [zero and 50/3]

16. Let X and Y be independent random variables having variances k and 2 respectively. If the variance of $3X - Y$ is 25, show that $k = 23/9$.

17. A horizontal line of length a units is divided into 2 parts. If the first part is of length X, find $E(X)$, Var (X) and $E[X(a-x)]$. [$\mu = a^2/3, \sigma^2 = a^2/12$ and $a^2/6$]

18. A drunk man with n keys wants to open his door and tries the keys independently and at random. Find the mean and variance of the number of trials required to open the door :
(i) If unsuccessful keys are not eliminated from further selection. [$\mu = n, \sigma^2 = n(n-1)$]
(ii) If they are eliminated. [[$\mu = (n+1)/2, \sigma^2 = (n^2 - 1)/12$].

19. A chord of a circle of radius a is drawn parallel to a given St. line, all distances from the centre of the circle being equally likely. Show that the expected value of the length of the chord is $\pi a/2$ and the variances of the length is $a^2(3^2(32) - 3\pi^2)/12$. Also, show that the chance is 1/2 that the length of the chord will exceed the length of the side of an equilateral triangle inscribed in the circle.

20. A large population consists of equal number of individuals of C different types. Individuals are drawn one at a time and put back before the next drawing. Denoting by N the smallest number of drawings which produce individual of each type, find $E(N)$ and Var (N).

5-90. *Expectation of Composite Functions*

When the function $Z = g(X, Y)$ whose expectation is sought is expressed in composite form, it is best to carry it alongwith its indicators to the elevation level. The following examples illustrate this modern way of evaluations. It is definitely more instructive and illuminating than the classical methods of evaluations presented earlier.

***Example* 1.** Let X be Expo (λ) and let $Y = \min(X, c)$. Find the Var (Y).

Solution. Here $Y = \min(X, c)$, so that $Y = X$ if $X \le c$, $Y = c$ if $X > c$. In terms of indicators $Y = XI_A + cI_B$. $I_A = \{0 < x \le c\}$. $I_B = \{X > c\}$. Thus

$$\mathrm{E}(Y) = \mathrm{E}(XI_A + cI_B) = \mathrm{E}(XI_A) + \mathrm{E}(cI_B). [f(x) = \lambda e^{-\lambda x}, x > 0]$$

$$= \int_0^c x . \lambda e^{-\lambda x} dx + c\int_0^\infty \lambda e^{-\lambda x} dx = [e^{-\lambda x}(x + \lambda^{-1})]_c^0 + c[e^{-\lambda x}]_\infty^c = \lambda^{-1}(1 - e^{-\lambda c})$$

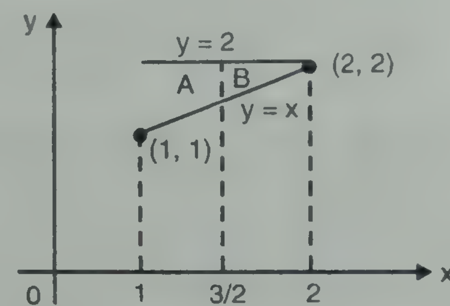
$$\mathrm{E}(Y^2) = \mathrm{E}(X^2 I_A + c^2 I_B) = \mathrm{E}(X^2 I_A) + \mathrm{E}(c^2 I_B) = \int_0^c x^2 \lambda e^{-\lambda x}\, dx + c^2 \int_c^\infty \lambda e^{-\lambda x} dx$$

$$= [e^{-\lambda x}(x^2 + 2x\lambda^{-1} + 2\lambda^{-2})]_c^0 + c^2[e^{-\lambda x}]_\infty^c = (2/\lambda^2)[1 - e^{-\lambda c}(1 + \lambda c)].$$

$$\therefore \qquad \mathrm{Var}(Y) = \lambda^{-2}[1 - 2\lambda c e^{-\lambda c} - 2e^{-\lambda c}]$$

***Example* 2.** Let $f(x, y) = 2$, on triangular region $1 \le x \le y \le 2$.
Compute $\mathrm{E}(X)$, $\mathrm{E}(Y)$, $\mathrm{E}(Z)$ where

$$Z = X \text{ for } X \le 3/2, \; Z = 2XY \text{ for } X > 3/2.$$



Solution. Here $f_1(x) = \int_x^2 f(x, y)\, dy = 2(2 - x), \; 1 \le x \le 2.$

$$\mathrm{E}(X) = \int_1^2 x f_1(x)\, dx = \int_1^2 (2x - x^2) dx = \left[x^2 - \frac{x^3}{3}\right]_1^2 = \frac{4}{3}.$$

$$\mathrm{E}(Y) = \int_1^2 \int_x^2 y f(x, y)\, dy\, dx = \int_1^2 (4 - x^2)\, dx = 4 - \frac{7}{3} = \frac{5}{3}.$$

$$\mathrm{E}(Z) = \mathrm{E}(XI_A + 2XYI_B) = \mathrm{E}(XI_A) + \mathrm{E}(XY\, I_B). \qquad [I_A = x \le 3/2,\; I_B = x > 3/2]$$

$$= \int_1^{3/2} 2x(2 - x)\, dx + 2\int_{3/2}^2 x \int_x^2 y . f(x, y)\, dy\, dx$$

$$= 2\int_1^{3/2} (2x - x^2)\, dx + 2\int_{3/2}^2 x(4 - x^2)\, dx$$

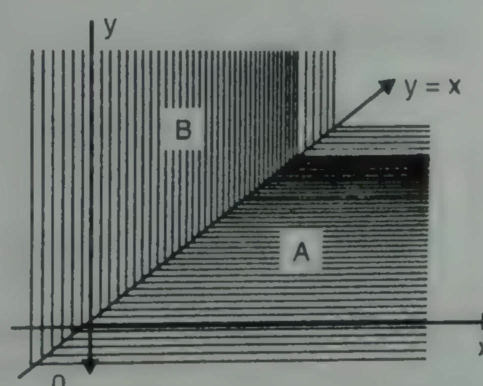
$$= 2[x^2 - (x^3/3)]_1^{3/2} + \tfrac{1}{2}[x^2(8 - x^2)]_{3/2}^2 = (11/12) + (49/32) = 235/96.$$

***Example* 3.** Let X and Y be independent variates with densities $f_1(x)$ and $f_2(y)$. Find $\mathrm{E}(Z)$ and Var (Z) when (a) $Z = |X - Y|$, (b) $Z = \min\{X, Y\}$. (c) $Z = \max\{X, Y\}$.

Solution. (a) Here $Z = X - Y$ if $X > Y$, and $Z = Y - X$ if $Y > X$.

Hence we write $Z = (X - Y)I_A + (Y - X)I_B$, where $I_A = \{(x, y) : x > y\}$, $I_B = \{(x, y) : y > x\}$

$$\therefore \mathrm{E}(Z) = \mathrm{E}[(X - Y)I_A] + \mathrm{E}[(Y - X)I_B] = J_1 + J_2 \text{ (say)} \quad ...(1)$$



$$J_1 = \iint_A (x - y)\, f_1(x)\, f_2(y)\, dx\, dy$$

$$= \int_{-\infty}^\infty x f_1(x)\left(\int_{-\infty}^x f_2(y)\, dy\right) dx - \int_{-\infty}^\infty y f_2(y)\left(\int_y^\infty f_1(x)\, dx\right)$$

$$= \int_{-\infty}^\infty x f_1(x)\, F_2(x)\, dx - \int_{-\infty}^\infty y f_2(y)[1 - F_1(y)]\, dy \quad ...(2)$$

$$J_2 = \int_{-\infty}^{\infty} y f_2(y)\, F_1(y)\, dy - \int_{-\infty}^{\infty} x f_1(x)\,[1 - F_2(x)]\, dx, \quad [x \rightleftharpoons y \text{ in } (2)] \qquad \ldots(3)$$

Substituting from (2) and (3) into (1) and collecting like terms, we obtain

$$\mathrm{E}(Z) = 2\int_{-\infty}^{\infty} x f_1(x) F_2(x)\, dx + 2\int_{-\infty}^{\infty} y f_2(y)\, F_1(y)\, dy - \mathrm{E}(X) - \mathrm{E}(Y) \qquad \ldots(4)$$

where $\mathrm{E}(X) = \int_{-\infty}^{\infty} x f_1(x)\, dx$ and $\mathrm{E}(Y) = \int_{-\infty}^{\infty} y f_2(y)\, dy$, are supposed known already.

Var (Z) can be had through Var (X) and Var(Y) instantly. Observe Var $(Z) = \mathrm{E}(Z^2) - \mathrm{E}^2(Z)$ ...(5)

$$\begin{aligned} \mathrm{E}(Z^2) &= \mathrm{E}(X - Y)^2 = \mathrm{E}(X^2) + \mathrm{E}(Y^2) - 2\mathrm{E}(XY) \\ &= [\mathrm{Var}(X) + \mathrm{E}^2(X)] + [\mathrm{Var}(Y) + \mathrm{E}^2(Y)] - 2\mathrm{E}(X)\, \mathrm{E}(Y) \qquad \ldots(6) \end{aligned}$$

From (4), (5) and (6), Var (Z) is obtained readily.

(b) Here $Z = X$, if $X \le Y$ and $Z = Y$, if $X > Y$. Hence we write

$Z = YI_A + XI_B$, where $I_A = \{(x, y) : x > y\}$, $I_B = \{(x, y) : x \le y\}$.

$$\mathrm{E}(Z) = \mathrm{E}(YI_A) + \mathrm{E}(XI_B) = J_1 + J_2, \text{ say.} \qquad \ldots(1)$$

$$J_1 = \iint_A y f_1(x)\, f_2(y)\, dx\, dy = \int_{-\infty}^{\infty} y f_2(y) \left(\int_y^{\infty} f_1(x)\, dx \right) dy = \int_{-\infty}^{\infty} y f_2(y)\,[1 - F_1(y)]\, dy \qquad \ldots(2)$$

$$J_2 = \iint_B x f_1(x)\, f_2(y)\, dx\, dy = \int_{-\infty}^{\infty} x f_1(x) \left(\int_x^{\infty} f_2(y)\, dy \right) dx = \int_{-\infty}^{\infty} x f_1(x)\,[1 - F_2(x)]\, dx \qquad \ldots(3)$$

Substituting from (2) and (3) into (1) we get

$$\mathrm{E}(Z) = \int_{-\infty}^{\infty} x f_1(x)\,[1 - F_2(x)]\, dx + \int_{-\infty}^{\infty} y f_2(y)\,[1 - F_1(y)]\, dy \qquad \ldots(4)$$

$$= \mathrm{E}(X) + \mathrm{E}(Y) - \int_{-\infty}^{\infty} x f_1(x)\, F_2(x)\, dx - \int_{-\infty}^{\infty} y f_2(y)\, F_1(y)\, dy. \qquad \ldots(5)$$

Since $Z^2 = Y^2 I_A + X^2 I_B$, we get through (4)

$$\mathrm{E}(Z^2) = \int_{-\infty}^{\infty} x^2 f_1(x)\,[1 - F_2(x)\, dx + \int_{-\infty}^{\infty} y^2 f_2(x)\,[1 - F_1(y)]\, dy \qquad \ldots(6)$$

and Var$(Z) = \mathrm{E}(Z^2) - \mathrm{E}^2(Z)$, is readily obtained through (4) and (6).

(c) Here $Z = X$ if $X > Y$ and $Z = Y$ if $Y > X$. Hence we write, as in (a) and (b) $Z = XI_A + YI_B$, where $I_A = \{(x, y) : x > y\}$, $I_B = \{(x, y) : x < y\}$

$$\mathrm{E}(Z) = \mathrm{E}(XI_A) + \mathrm{E}(YI_B) = J_1 + J_2, \text{ say} \qquad \ldots(1)$$

$$J_1 = \iint_A x f_1(x)\, f_2(y)\, dx\, dy = \int_{-\infty}^{\infty} x f_1(x) \left(\int_{-\infty}^{x} f_2(y)\, dy \right) dx = \int_{-\infty}^{\infty} x f_1(x)\, F_2(x)\, dx \qquad \ldots(2)$$

$$J_2 = \iint_B y f_1(x)\, f_2(y)\, dx\, dy = \int_{-\infty}^{\infty} y f_2(y) \left(\int_{-\infty}^{y} f_1(x)\, dx \right) dy = \int_{-\infty}^{\infty} y f_2(y)\, F_1(y)\, dy \qquad \ldots(3)$$

Substituting from (2) and (3) into (1) we get

$$\mathrm{E}(Z) = \int_{-\infty}^{\infty} x f_1(x)\, F_2(x)\, dx + \int_{-\infty}^{\infty} y f_2(y)\, F_1(y)\, dy. \qquad \ldots(4)$$

Since, $Z^2 = X^2 I_A + Y^2 I_B$, equation (4) yields

$$E(Z^2) = \int_{-\infty}^{\infty} x^2 f_1(x)\, F_2(x)\, dx + \int_{-\infty}^{\infty} y^2 f_2(y)\, F_1(y)\, dy. \qquad \ldots(5)$$

$\text{Var}(Z) = E(Z^2) - E^2(Z)$, is readily obtained through (4) and (5).

Range Z is obtainable by 3σ rule : $\mu_Z \pm 3\sigma_Z$.

Exercises 5(d)

1. Let $f(x, y) = 1/2$, on the square with vertices (0, 1), (1, 0), (2, 1) and (1, 2). Let $M = \max(X, Y)$ and define $Z = X$ if $M \le 1$, $Z = 2Y$ if $M > 1$. Prove that $\text{Var}(Z) = 67/72$.
2. Let $f(x, y) = 3$ for $0 \le y \le x_2 \le 1$. Let $Z = X$ if $X + Y \le 1$; $Z = 1$ if $X + Y > 1$. Show that $E(X) = 3/4$, $E(Y) = 3/10$, $E(Z) = 3(5\sqrt{3} - 9)/8$.
3. Ram plans to stock n units of Daffodill Eye-liners, at a cost of c per unit, Demand D ~ Pois (λ). If units remain unsold, these may be returned for r per unit. If demand exceeds n, the extra units can be had at a cost e per unit. Show that the expected gain, if the units are sold at a price p each, is
   $E(X) = \lambda(p - r) - \lambda(e - r)\, P(D \ge n) + n(e - r)\, P(D \ge n + 1) - n(c - r)$.
4. Let $X \sim N(\mu, \sigma^2)$ and let $Y = \min(X, c)$, find $\text{Var}(Y)$.
5. Find $E(Y)$ and $\text{Var}(Y)$ if $Y = |X|$ and X has density $f_X(x)$.
6. Let $X \sim$ unif $(-2, 2)$. Define r.v. Z by $Z = X$ if $X < 1$, $Z = 1$ if $X \ge 1$
   Find distribution of Z and its variance.

One cloud is enough to eclipse all the sun.

Praise undeserved is satire in disguise. (Alex Pope)

Moments and Quantiles 6

6-10. *Moments of Random Variables*

The rth order moment of a random variable X about a constant a is denoted by μ'_r and is defined by

$$\begin{aligned}\mu'_r &= \mathsf{E}[(X-a)^r] = \int (x-a)^r \, dF(x), \quad \text{(general)}\\ &= \int_{-\infty}^{\infty} (x-a)^r f(x)\,dx, \quad \text{when p.d.f. exists}\\ &= \Sigma_i (x_i - a)^r \, p(x_i), \quad \text{when } X \text{ is discrete.}\end{aligned}$$

We recall that, $\mathsf{E}[(X-a)^r]$ exists iff $\mathsf{E}[|(X-a)^r|]$ exists. When $a = 0$, then $\mathsf{E}(X^r)$ is called the rth order *simple moment* and in particular, $\mathsf{E}(X) = \mu'_1 = \mu$ (say) is called the *centre* on (arithmetic) *mean* of the distribution.

The moments about the point $\mathsf{E}(X)$ are called *central moments* and central moment of order r is denoted by μ_r and is defined by

$$\mu_r = \mathsf{E}(X-\mu)^r \;\; ; \;\; \mu = \mathsf{E}(X) = \mu'_1.$$

Simple moments are also termed *raw* moments or *ordinary* moments.

6-11. *Moments : Higher-order versus Lower-order*

If $\mathsf{E}(X^r)$ exists, then $\mathsf{E}(X^k)$ necessarily exists, where $0 < k < r$.

Proof. (i) Let X be a continuous r.v. with density $f(x)$.

Since the rth order moment is given to exist, the integral $\int_{-\infty}^{\infty} x^r f(x)\,dx$ converges absolutely, i.e. $\int_{-\infty}^{\infty} |x|^r f(x)\,dx < \infty$. Now for every $k = 1, 2, \ldots, r-1$, we have $|x|^k < |x|^r + 1$. Consequently

$$\int_{-\infty}^{\infty} |x|^k f(x)\,dx < \int_{-\infty}^{\infty} |x|^r f(x)\,dx + 1 < \infty.$$

(ii) When X is a discrete variate, we observe that

$$|x| \le 1 \;\Rightarrow\; |x|^k \le 1; |x| > 1 \;\Rightarrow |x|^k \le |x|^r, \;\Rightarrow\; |x|^k < |x|^r + 1.$$

By Comparison theorem for convergent series, (i) provides

$$\Sigma_x |x|^k f(x) < \Sigma_x [|x|^r + 1] f(x) \text{ i.e. } \mathsf{E}(|X|^k) < \mathsf{E}(|X|^k) + 1 < \infty. \qquad [\Sigma_x f(x) = 1]$$

Thus $\mathsf{E}(X^k)$ certainly exists.

***Example* 1.** If $E(X^r)$ exists, then $E(X^t)$ need not exist if $t > r$.

Solution. Let $\qquad f(x) = (r+1)\lambda^{r+1}/(x+\lambda)^{r+2}, x \geq 0, \lambda > 0.$ ...(1)

$$E(X^r) = (r+1)\lambda^{r+1}\int_0^\infty \frac{x^r\,dx}{(x+\lambda)^{r+2}} = (r+1)\lambda^r \int_0^\infty \frac{z^r\,dz}{(1+z)^{r+2}} = (r+1)\lambda^r B(r+1,1) = \lambda^r \quad [x = \lambda z]$$

Thus, $E(X^r) = \lambda^r$ (finite), i.e. μ'_r exists.

However : $E(X^{r+1}) = (r+1)\lambda^{r+1}\int_0^\infty \frac{x^{r+1}dx}{(x+\lambda)^{r+2}} \to \infty.$

because the later integral is divergent. Hence for the p.d.f. (1) only moments upto order r exist, all higher order moments do not exist.

Note. Some distributions which don't possess moment of any order are given in examples 5-1 and 5-2.

6-12. *Moment Problem*

A c.d.f. determines a set of moments when they exist. However, the determination of c.d.f. 'F', when all necessary moments exist is subject to varying restrictions. Some of the results are :

1. Let F be a c.d.f. of a variate X, $E(X^r) = \mu'_r < \infty$, $\forall\ r = 1, 2, 3, \ldots$. If the infinite series $\Sigma \mu'_r t^r / r!$, $r = 1, 2, \ldots$ is obsolutely convergent for some $t > 0$, then $F(x)$ is the unique c.d.f. possessing these moments.
2. For a bounded variate $[P(|X| < B) = 1]$, its c.d.f. is uniquely determined by its moments.
3. *Carleman's Criterion* : Let $E(X^r) = \mu'_r < \infty$ for $r = 1, 2, 3, \ldots,\quad -\infty < x < \infty.$

If $\Sigma[(\mu'_{2r})^{-1/2r}] = 1 \leq r < \infty$, then the sequence $\{\mu'_r\}$ uniquely determines the c.d.f. of X.

4. Some *sufficient conditions* for a moment sequence to determine a unique c.d.f. are :

(i) The range of the variate is finite. (ii) $\Sigma(1/\mu'_r)^{1/2r} = \infty, r = 0, 1, 2, \ldots \infty,\ 0 \leq x < \infty.$

(iii) $\limsup\limits_{n\to\infty}\left[\frac{(\mu'_{2n})^{1/2n}}{2n}\right] < \infty.$

6-13. *Worked-out Problems*

***Example* 1.** Give an example of a distribution which is not uniquely determined by its moments.

Solution. Consider ℓn-Normal distribution : $f(x) = (x\sqrt{2\pi})^{-1}\exp\left[-\frac{1}{2}(\ln x)^2\right], x > 0;\ f(x) = 0, x \leq 0.$

Let $\qquad g(x) = [1 + a\sin(2\pi \ln x)]\, f(x) = f_a(x)$, say. $\qquad [-1 \leq a \leq 1]$

$\therefore \qquad \int_0^\infty x^r g(x)\,dx = \int_0^\infty f(x)\,x^r\,dx + a\int_0^\infty x^r \sin(2\pi \ln x)\, f(x)\,dx$ ...(1)

Let $\qquad J = \int_0^\infty x^n \left(x^{-1} e^{-(1/2)[\ln x]^2}\right)\sin(2\pi \ln x)\,dx$

Put $\ell n\, x = u = y + n$; then

$$J = \int_{-\infty}^{\infty} e^{-(1/2)u^2 + nu} \sin(2\pi u)\, du = e^{(1/2)n^2} \int_{-\infty}^{\infty} e^{-(1/2)y^2} \sin(2\pi y)\, dy.$$

The last integrated is an odd function of y, hence $J = 0$, whence (1) provides

$$\int_0^{\infty} x^r\, g(x)\, dx \;=\; \int_0^{\infty} x^r\, f(x)\, d.$$

Thus, $[f_a : -1 \le a \le 1\}$ is a collection of density functions each different from all the others, but all having the same moments.

Comments. This example shows that the moments may not uniquely define a density function.

Example 2. Let $\alpha > 0$ be any constant. If a r.v. X is such that $n^{\alpha} P(|X| > n) \to 0$ as $n \to \infty$, then show that $\mathsf{E}(|X|^{\beta}) < \infty$ for $0 < \beta < \alpha$.

Solution. We use Def. of limit. For brevity, set $T_n = P\{|X| > n\}$. Since $n^{\alpha} T_n \to 0$ as $n \to \infty$, so given $\varepsilon > 0$, there is $n_0\ (\varepsilon) = n_0$ (say), such that

$$n^{\alpha} T_n < \varepsilon,\ \forall\ n \ge n_0, \text{ i.e. } T_n < \varepsilon / n^{\alpha},\ \forall\ n \ge n_0. \qquad \ldots(1)$$

Recall geometric concept : $\mathsf{E}(Y) = \int_0^{\infty} P(Y > t)\, dt$ [§ 5-11]. For $0 < \beta < \alpha$, we have

$$\mathsf{E}(|X|^{\beta}) = \int_0^{\infty} P(|X|^{\beta} \ge x)\, dx = \beta \int_0^{\infty} x^{\beta - 1} P(|X| > x)\, dx \; [x = z^B \text{ ; then change dummy } z \text{ to } x]$$

$$= \beta \int_0^{n_0} x^{\beta - 1} T_x\, dx + \beta \int_{n_0}^{\infty} x^{\beta - 1} T_x\, dx$$

$$\le \beta \int_0^{n_0} x^{\beta - 1} dx + \beta\, \varepsilon \int_{n_0}^{\infty} x^{\beta - \alpha - 1} dx \; [T_x < 1 \text{ in first term, } T_x < \varepsilon / x^d \text{ in 2nd term}] \quad \ldots(2)$$

$$= (n_0)^{\beta} - [\beta \varepsilon / (\alpha - \beta) n_0^{\alpha - \beta}] \quad \text{(finite quantity)}$$

Thus $\mathsf{E}(|X|)^{\beta} < \infty$, i.e. absolute moment of order β esists.

Problems with Solutions Provided at the End of the Text

1*. Show that if r.v. X is bounded, it has moments of every order.

2*. Let X be a r.v. such that $\mu'_k < \infty$, for some $k > 0$. Show that $\lim_{n \to \infty} n^k P\{|X| > n\} = 0$.

Does the converse hold ?

Exercise 6(a)

1. For a random variable X with p.d.f.

$$f(x) = \tfrac{1}{2} x,\ 0 \le x < 1;\ f(x) = \tfrac{1}{2},\ 1 < x \le 2;\ f(x) = \tfrac{1}{2}(3 - x),\ 2 < x \le 3;$$

show that moments of all orders exist.

2. Let X be a r.v. with p.d.f. : $f(x) = k[1 + (x^2/a^2)]^{-m} \exp[-\lambda \tan^{-1}(x/a)],\ -\infty < x < \infty.$

Show that : $(2m - r - 1)\ \mu'_r = a[(r - 1)\, a\mu'_{r-2} - \lambda \mu'_{r-1}]$.

3. Show that for the distribution : $dF = k[1 - (x^2/a^2)]^{-p}\, dx,\ 0 < p < 1,\ -a \le x <= a$

$$\mu_{r+1} = [ra^2 / (r + 2 - 2p)]\, \mu_{r-1}$$

Deduce that $\dfrac{a^{2r}}{a^{2s}} < \dfrac{2r+1}{2s+1}\dfrac{\mu_{2r}}{\mu_{2s}}$, $r > s$.

4. Let X be a continuous variate with d.f. $F(x)$. Show that $\mu_r' = \int_0^\infty r x^{r-1} [1 - F(x) + (-1)^r F(-x)]\, dx$.

   Formulate a similar result if X is a discrete variate taking the non-negative integral values a, $a + 1, \ldots, b$ (a finite).

5. Prove the moment lemma

$$E\{|X|^\alpha\} < \infty \Leftrightarrow \Sigma P\{|X| > n^{1/\alpha}\} < \infty, [n = 1, 2, \ldots]$$

6. Let X be a variate satisfying $\{P\{\,|X| > \alpha k\}/P\{\,|X| > k\}\} \to 0$, as $k \to \infty$ for all $\alpha > 1$. Show that X posseses moment of all orders.

7. If $f(x)$ is an odd function of x with period $\frac{1}{2}$, show that $\int_0^\infty x^r x^{-\ln x} f(\ln x)\, dx = 0$ for all integral values of r. Hence show that the distribution $dF(x) = x^{-\ln x}\, [1 - \lambda \sin (4\pi \ln x)]\, dx$, $0 < x < \infty$, $0 \le \lambda \le 1$ have the same moments whatever the value of λ.

8. For the variates X and Y, let $E(X^r) = \mu_r'$, $E(Y^r) = m_r'$. Show that the rth raw (simple) moment of $X + Y$ is $E(X + Y)^r = \sum_{s=0}^{r} \binom{r}{s} m_{r-s}' \ldots$

9. If $F(x) = (1 - e^{-ax})\, u(x - a)$, find $f(x)$ and $E(X^n)$, where $u(x - a)$ is a unit step function.

10. Let X and Y be continuous variates with densities 'f' and g. If $E(X^k) = E(Y^k)$, $1 \le k \le n$, show that the graphs of $z = f(x)$ and $z = g(x)$ cross each other at atleast $(n + 1)$ points.

6-20. *Relation between Moments about any Two Points*

Let $X - a = X_a$, $X - b = X_b$, put $b - a = c$; then $X_a = X_b + c$. Now using Binomial expansion and lin E, we get

$$E(X - a)^n = E(X_b + c)^n = E\left\{\sum_{r=0}^{n} \binom{n}{r} (X_b)^{n-r} c^r\right\} = \sum_{r=0}^{n} \binom{n}{r} E(X_b)^{n-r} c^r.$$

$$\boxed{E(X - a)^n = \sum_{r=0}^{n} \binom{n}{r} (b - a)^r\, E(X - b)^{n-r}}$$

Cor. Relation between central and non-central moments

1. $E(X - \mu)^n = \sum_{r=0}^{n} \binom{n}{r} E(X - b)^{n-r} . (-1)^r (\mu - b)^r \Rightarrow \mu_n = \sum_{r=0}^{n} \binom{n}{r} \mu_{n-r}' (-\mu_1')^r$. $[\mu_1' = \mu - b]$

2. $E(X - a)^n = \sum_{r=0}^{n} \binom{n}{r} E(X - \mu)^{n-r} . (\mu - a)^r \Rightarrow \mu_n' = \sum_{r=0}^{n} \binom{n}{r} \mu_{n-r} (\mu_1')^r$. $[\mu_1' = \mu - a]$

First four particular values. $[\mu_0 = 1, \mu_1 = 0]$

$$\boxed{\mu_2 = \mu_2' - (\mu_1')^2, \quad \mu_3 = \mu_3' - 3\mu_2'\mu_1' + 2(\mu_1')^3, \quad \mu_4 = \mu_4' - 4\mu_3'\mu_1' + 6\mu_2'(\mu_1')^2 - 3(\mu_1')^4}$$

$$\mu_2' = \mu_2 + \mu_1'^2, \mu_3' = \mu_3 + 3\mu_2\mu_1' + (\mu_1')^3, \quad \mu_4' = \mu_4 + 4\mu_3\mu_1' + 6\mu_2(\mu_1')^2 + (\mu_1')^4.$$

6-21. *Effect of Change of Origin and Scale on Moments*

By definition, $\mu_n = E(X - \mu)^n$; hence if we put $(X - a)/h = U$, then $X = a + hU$,

$$E(X) = a + hE(U) \Rightarrow \mu_n = E\{(a + hU) - [a + hE(U)]\}^n = h^n E[\{U - \mu_U\}^n]$$

Thus, the nth central moment is invariant for the change of origin. However, the nth central moment of X is h^n times the nth central moment of variate U.

6-22. *Standardization (or Normalization)*

A r.v. Z is said to be standard (or standardized or reduced) if $E(Z) = 0$ and var $(Z) = 1$. To standardize a non-standard variate X, we let $\mu = E(X)$, $\sigma = \sqrt{Var(X)}$ and construct a variate

$$X^* = (X - \mu)/\sigma. \qquad ...(1)$$

Obviously, $E(X^*) = 0$, Var $(X^*) = 1$, hence X^* is a standardized variate. Eqn. (1) gives the method of standardizing a non-standard variate X.

Moments of a Standardized Variate

Let $$Z = [X - E(X)]/\sqrt{Var(X)} = (X - \mu)/\sigma$$

The nth-order moment of Z is denoted by α_n and is defined by

$$\alpha_n = E(Z^n) = E[(X - \mu)^n/\sigma^n] = E(X - \mu)^n/\sigma^n = \mu_n/\sigma^n$$

Thus $$\boxed{\alpha_n = \mu_n/\sigma^n, \quad \alpha_1 = 0, \quad \alpha_2 = 1, \quad \alpha_3 = \mu_3/\sigma^3, \quad \alpha_4 = \mu_4/\sigma^4, \ldots}$$

Some authors use the notation of alpha, beta and gamma coefficients :

$$\beta_1 \equiv \alpha_3^2, \quad \beta_2 \equiv \alpha_4, \quad \gamma_1 = \alpha_3, \quad \gamma_2 = \alpha_4 - 3,$$

$$\beta_{2r} = \alpha_{2r+2} = \mu_{2r+2}/(\sigma)^{2r+2}, \qquad \beta_{2r+1} = (\alpha_{2r+3})(\alpha_3) = \mu_{2r+3} \cdot \mu_3/\sigma^{2r+6}.$$

6-23. *Pearson's Shape Coefficients*

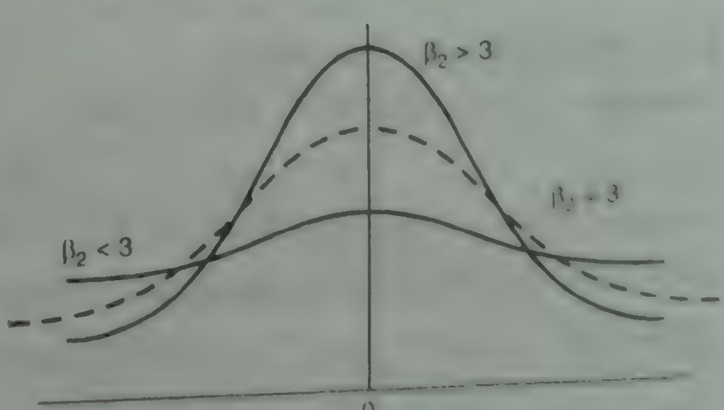
The dimensionless coefficients β_1, β_2, γ_1, γ_2 are called Pearson's coefficients.

The graph of a distribution is said to be

(i) Platykurtic, if $\beta_2 < 3$.

(ii) Mesokurtic (or normal), if $\beta_2 = 3$.

(iii) Leptokurtic, if $\beta_2 > 3$.



Note. Platykurtic (Broad) curves have short tails like a platypus while leptokurtic (Thin) curves have long tails like kangaroos noted for lepping.

Peakiness. β_2 is called *kurtosis* (or peakedness) of the probability distribution and can be regarded as a measure of its degree of 'peakedness'. Also, min $\beta_2 = 1$, this value is attained by a variate which takes only two values with equal probabilities.

Coefficient of Skewness. The constant $\gamma_1 = \sqrt{\beta_1}$ is called the *coefficient of skewness* (S_k) and measures departure from symmetry of a frequency curve. It is invariant under linear transformation. If $\gamma_1 = 0$, the curve is not skewed, if $\gamma_1 > 0$, the curve is positively skewed ; and if $\gamma_1 < 0$, the curve is negatively skewed.

These are important dimensionless quantities described earlier.

In the normal distribution, to be considered later, $\beta_2 = 3$.

Since normal distribution is regarded as the standard or the ideal distribution, the quantity $\beta_2 - 3 = \gamma_2$, for any distribution is called its *excess of kurtosis*, or briefly its *excess*.

Relation between Pearson's Moments Coefficients

For any variate X, for which β_1 and β_2 exist, and for any k,

$$\beta_2 \geq \beta_1 - (2k + k^2), \qquad \beta_2 \geq \beta_1, \qquad \beta_2 \geq \beta_1 + 1.$$

Proof. Let $Y = X - \mathrm{E}(X)$; and consider the **non-negative** quadratic form

$$Q(t) = \mathrm{E}(Y^2 + tY + k\mu_2)^2 = \mathrm{E}[Y^4 + 2tY^3 + (t^2 + 2k\mu_2)Y^2 + k^2\mu_2^2 + 2t\mu_2 kY] \geq 0$$

$$= t^2\mu_2 + 2t\mu_3 + [\mu_4 + (2k + k^2)\mu_2^2] \geq 0. \qquad \text{[by Lin E]} \qquad \ldots(1)$$

Since $Q(t) \geq 0$, its discriminant is negative, hence

$$\mu_3^2 - \mu_2[\mu_4 + (2k + k^2)\mu_2^2] \leq 0; \text{i.e.} \ \mu_4 - (\mu_3^2/\mu_2) + (k^2 + 2k)\mu_2^2 \geq 0 \Rightarrow \beta_2 \geq \beta_1 - (2k + k^2).$$

If $k = 0, -1$, we get $\beta_2 \geq \beta_1$; $\beta_2 \geq \beta_1 + 1$. ...(2)

Remarks. We could differentiate (1) to get $Q'(t) = 0$ giving $t = -\mu_3/\mu_2$. The result (1) is true for all values of t ; in particular for $t = -\mu_3/\mu_2$. And with this value, (1) reduces to (2).

General Relations. $\beta_{2r-2} > \beta_{2r-3} + \beta_{2r-4}$

In particular : $\beta_{2r-2} > 1 + \beta_1 + \beta_2 + \ldots + \beta_{2r-3}$.

6-25. *Worked-out Problems*

Example : Show that, if β_2 exists, then

(i) $\beta_2 \geq \beta_1$ (ii) $\beta_2 \geq 1$ (iii) When is $\beta_2 = 1$?

Solution. Let $Y = X - \mu_X$, then $\mathrm{E}(Y) = 0$, $\mathrm{E}(Y^r) = \mathrm{E}(X - \mu_X)^r = \mu_r$. Consider $Q(\lambda) = \mathrm{E}(Y^m + \lambda Y^n)^2$ Then, $Q(\lambda) \geq 0$ for *all* real values of λ. Now

$$Q(\lambda) = \mathrm{E}(Y^{2m} + 2\lambda Y^{m+n} + \lambda^2 Y^{2n}) = \mathrm{E}(Y^{2m}) + 2\lambda \mathrm{E}(Y^{m+n}) + \lambda^2 \mathrm{E}(Y^{2n}) = \lambda^2\mu_{2n} + 2\lambda\mu_{m+n} + \mu_{2m} \geq 0$$

Since $Q(\lambda) \geq 0$, $\forall\ \lambda \in R$, its discriminant is negative, whence $\mu_{2m}\mu_{2n} \geq (\mu_{m+n})^2$. ...(1)

(i) Choose $m = 2$, $n = 1$, then (1) gives $\mu_4\mu_2 \geq \mu_3^2$ or $(\mu_4/\mu_2^2) \geq (\mu_3^2/\mu_2^3) \Rightarrow \beta_2 \geq \beta_1$.

(ii) Choose $m = 2$, $n = 0$, then (1) provides $\mu_4 \geq \mu_2^2$ or $(\mu_4/\mu_2^2) \geq 1 \Rightarrow \beta_2 \geq 1$.

(iii) $\beta_2 = 1 \Rightarrow \mu_4 - \mu_2^2 = 0 \Rightarrow \mathrm{E}(Y^2 - \mu_2)^2 = \text{Var}\, Y^2 = 0$. Thus

$$P\{Y^2 = \mathrm{E}(Y^2)\} = 1, \quad \text{i.e. } P\{Y = \pm\sigma\} = 1.$$

Let $P\{Y = +\sigma\} = p, P\{Y = -\sigma\} = q$ and $\mathrm{E}(Y) = 0$; then $\mathrm{E}(Y) = \sigma p + (-\sigma)q = \sigma(p - q)$.

Now $\sigma \neq 0$ (otherwise β_2 is not defined), hence $p = q \Rightarrow p = \frac{1}{2} = q$.

We conclude that $\beta_2 = 1$ iff variate X assumes only two values with equal probabilties, each $\frac{1}{2}$.

Problems with Solutions Provided at the End of the Text

1*. The first three moments about origin are

$$\mu_1' = \tfrac{1}{2}(n+1),\ \mu_2' = \tfrac{1}{6}(n+1)(2n+1),\ \mu_3' = \tfrac{1}{4}n(n+1)^2.$$

Determine the skewness of the data.

2*. Let $f(x) = [2\Gamma(1+\theta)]^{-1}\lambda^{\theta}\exp[-\lambda|x|^{\theta}];\ \theta > 0,\ \lambda > 0;\ -\infty < x < \infty.$

Show that $\beta_2 = \Gamma(5\theta)\Gamma(\theta)/[\Gamma(3\theta)]^2$.

3*. The first three moments of a distribution about the value 2 of the variate X are 1, 16, -40 respectively. Show that the mean is 3, $\sigma^2 = 15$ and $\mu_3 = -86$. Find also the first three simple moments.

4*. The first four moments of a distribution about $X = 4$ are 1, 4 , 10, 45 respectively. Show that the mean is 5, $\sigma^2 = 3$, $\mu_3 = 0$ and $\mu_4 = 26$. Determine β_1 and β_2.

6-26. *Absolute Moments*

The rth order absolute moments of a r.v. X about $X = a$ is denoted by ν_r and is defined by

$$\nu_r = \mathrm{E}(|X - a|^r) = \mathrm{E}(|Y|^r). \quad [Y = X - a]$$

Some relations : (i) $\nu_r^2 \le \nu_{r-1}\cdot\nu_{r+1}$ (ii) $(\nu_r)^{1/r} \le (\nu_{r+1})^{1/(r+1)}$.

Proof. Consider the non-negative quantity Q defined as under :

$$Q = \mathrm{E}[\lambda|Y|^m + |Y|^n]^2 = \mathrm{E}[\lambda^2|Y|^{2m} + 2\lambda|Y|^{m+n} + |Y|^{2n}] = \lambda^2\,\mathrm{E}|Y|^{2m} + 2\lambda\,\mathrm{E}|Y|^{m+n} + \mathrm{E}|Y|^{2n}$$

i.e. $$Q = \lambda^2\nu_{2m} + 2\lambda\nu_{m+n} + \nu_{2n} \ge 0 \qquad \text{...(1)}$$

Since $Q \ge 0$, its discriminant is non-positive $\Rightarrow (\nu_{m+n})^2 \le \nu_{2m}\,\nu_{2n}$

(i) Choose $2m = r - 1$, $2n = r + 1$; the result (1) follows, viz $\nu_r^2 \le \nu_{r-1}\cdot\nu_{r+1}$. ...(2)

(ii) Replace r by k in (2), and raise it to power k to get $\nu_k^{2k} \le (\nu_{k-1})^k(\nu_{k+1})^k$. We rewrite it as

$$(\nu_k/\nu_{k-1})^k\,(\nu_{k+1}/\nu_k)^k. \qquad \text{...(3)}$$

Put $k = 1, 2, 3, \ldots, r-1, r$ in (3) ; then multiply these inequalities to obtain

$$\left(\frac{\nu_1}{\nu_0}\right)\cdot\left(\frac{\nu_2}{\nu_1}\right)^2\cdot\left(\frac{\nu_3}{\nu_2}\right)^3\cdots\left(\frac{\nu_{r-1}}{\nu_{r-1}}\right)^{r-1}\left(\frac{\nu_r}{\nu_{r-1}}\right)^r \le \frac{\nu_2}{\nu_1}\left(\frac{\nu_3}{\nu_2}\right)^2\cdots\left(\frac{\nu_r}{\nu_{r-1}}\right)^{r-1}\cdot\left(\frac{\nu_{r+1}}{\nu_r}\right)^r$$

Using $\nu_0 = 1$, this can be simplified to

$$\frac{1}{\nu_1 \nu_2} \cdots \frac{1}{\nu_{r-1}} (\nu_r)^r \le \frac{1}{\nu_1 \nu_2} \cdots \frac{1}{\nu_{r-1}} \frac{1}{\nu_r} (\nu_{r+1})^r$$

or
$$(\nu_r)^{r+1} \le (\nu_{r+1})^r \Rightarrow (\nu_r)^{1/r} \le (\nu_{r+1})^{1/(r+1)}$$

Note. Replacing a by μ_X and disregarding modulus sign, Eq. (1) gives $\mu^2_{(m+n)} \le \mu_{2m} \cdot \mu_{2n}$.

This then solves the problem : $\beta_2 \ge \beta_1$ and $\beta_2 \ge 1$, etc.

6-30. *Factorial Moments*

The *descending* factorial power with span-unit h is defined by

$$x^{(r)} = x(x-h)(x-2h)\ldots[x-(r-1)h] \quad \ldots(1)$$

The rth order descending factorial moment about the point $X = a$ is defined by

$$\mu'_{(r)} = \mathsf{E}[(X-a)^{(r)}] = \Sigma f(x)(x-a)^{(r)}. \quad \text{[Sum over } x\text{]} \quad \ldots(2)$$

Since
$$(a+b)^{(n)} = \Sigma \binom{n}{k} a^{(n-k)} b^{(k)}, \qquad \text{[Factorial Bin. Expansion]}$$

$$\therefore \quad (X-a)^{(n)} = [(X-b)+c]^{(n)} = \Sigma \binom{n}{k} (X-b)^{(n-k)} x^{(k)}, \quad (c = b-a).$$

Taking expectation of both sides,we obtain

$$\mu'_{(n)} = \Sigma \binom{n}{k} c^{(k)} \mu''_{(n-k)}, 0 \le k \le n. \quad [\mu''_{(r)} = \mathsf{E}(X-b)^{(r)}] \quad \ldots(3)$$

Often this result is written symbolically as

$$\mu'_{(n)} = (\mu'' + c)^{(n)}. \quad \ldots(4)$$

Remarks. In practice, we take the span-unit $h = 1$.

Some Relations between Simple and Factorial Moments

By direct expansion of (2) in §6-30, using $h = 1$, $a = 0$, we get

$$\mu'_{(1)} = \mathsf{E}(X) = \mu'_1.$$

$$\mu'_{(2)} = \mathsf{E}[X(X-1)] = \mathsf{E}(X^2) - \mathsf{E}(X) = \mu'_2 - \mu'_1.$$

$$\mu'_{(3)} = \mathsf{E}(X(X-1)(X-2)] = \mathsf{E}(X^3 - 3X^2 + 2X) = \mu'_3 - 3\mu'_2 + 2\mu'_1.$$

$$\mu'_{(4)} = \mathsf{E}(X(X-1)(X-2)(X-3)] = \mathsf{E}(X^4 - 6X^3 + 11X^2 - 6X) = \mu'_4 - 6\mu'_3 + 11\mu'_2 - 6\mu'_2.$$

Converse relations

$$\mu'_2 = \mathsf{E}(X^2) = \mathsf{E}(X^{(2)} + X) = \mu'_{(2)} + \mu'_{(1)}$$

$$\mu'_3 = \mathsf{E}(X^3) = \mathsf{E}(X^{(3)} + 3X^{(2)} + X) = \mu'_{(3)} + 3\mu'_{(2)} + \mu'_{(1)}$$

$$\mu'_4 = \mathsf{E}(X^4) = \mathsf{E}(X^{(4)} + 6X^{(3)} + 7X^{(2)} + X) = \mu'_{(4)} + 6\mu'_{(3)} + 7\mu'_{(2)} + \mu'_{(1)}.$$

6-31. *Central Factorial Moments in terms of Central Moments*

For brevity let $Y = X - \mathsf{E}(X)$ and note that $\mathsf{E}(Y) = 0$.

$$\mu_{(2)} = \mathsf{E}[Y^{(2)}] = \mathsf{E}[Y(Y-1)] = \mathsf{E}(Y^2) - \mathsf{E}(Y) \Rightarrow \mu_{(2)} = \mu_2$$

$$\mu_{(3)} = \mathsf{E}[Y^{(3)}] = \mathsf{E}[Y(Y-1)(Y-2)] = \mathsf{E}(Y^3 - 3Y^2 + 2Y) = \mathsf{E}(Y^3) - 3\mathsf{E}(Y^2) \Rightarrow \mu_{(3)} = \mu_3 - 3\mu_2$$

$$\mu_{(4)} = \mathsf{E}[Y^{(4)}] = \mathsf{E}[Y(Y-1)(Y-2)(Y-3)] = \mathsf{E}[Y^4 - 6Y^3 + 11Y^2 - 6Y] = \mathsf{E}(Y^4) - 6\mathsf{E}(Y^3) + 11\mathsf{E}(Y^2)$$

$$\therefore \quad \mu_{(4)} = \mu_4 - 6\mu_3 + 11\mu_2.$$

6-32. *Reverse or Increasing Factorial Moments*

Write : $X^{[k]} = X(X+1)(X+2)\ldots(X+k-1) = k!\binom{X}{k}$ Pichhommer symbols (notation)

Define : $\mu'_{[k]} = \mathsf{E}[X^{[k]}] = \mathsf{E}[X(X+1)\ldots(X+k-1)]$

This is kth order simple *reverse* factorial moment; at times, it will prove more useful than $\mu'_{[k]}$.

Exercise 6(b)

1. Show that first n simple moments determine the first n central moments and conversely.

$$[\mu_n = \mathsf{E}(X-\mu)^n = \mathsf{E}\left[\sum_{k=0}^{n}(-1)^{n-k}\binom{n}{k}X^k\mu^{n-k}\right] = \sum_{k=0}^{n}(-1)^{n-k}\binom{n}{k}\mu'_k\,\mu^{n-k}$$

$$\mu'_n = \mathsf{E}[(X-\mu)+\mu]^n = \mathsf{E}\left[\sum_{k=0}^{n}\binom{n}{k}\mu^{n-k}(X-\mu)^k\right] = \sum_{k=0}^{n}\binom{n}{k}\mu_k\,\mu^{n-k}\,]$$

2. The first three moments of a distribution about the value 3 of the variate X are 2, 10, – 30 respectively. Show that the first three moments about $X = 0$ are 3, 31, 141. Find also μ_2, μ_3.
3. The first four moments of a distribution about the value 5 of the variable X are 2, 20, 40, 50. Obtain, as far as possible, the various characteristics of the distribution on the basis of given information. [**Ans.** $\mu = 7, \mu_2 = 16, \mu_3 = -64, \mu_4 = 162, \beta_2 = 1, \beta_2 = 0.63$]
4. The first four moments of a distribution about the value 4 of the variable X are –1.5, 17, –30 and 108. Calculate the simple and central moments and state whether the distribution is leptokurtic or playtycurtic. Can you find moments about $X = 2$?
[**Ans.** $\mu_2 = 14.75$; $\mu_3 = 39.75$, $\mu_4 = 142.3$]
5. The distribution of variate X has mean 10, variance 16, $\gamma_1 = 1$ and $\beta_2 = 4$. Obtain the first four simple moments. [**Ans.** 10, 116, 1544, 23184]
6. In calculating the moments of a frequency distribution based on 100 observations, the following results were obtained : Mean = 9, Variance = 19, $\beta_1 = 0.7$, $\beta_2 = 4$.
But later on, it was found that one observation 12 was read as 21. Obtain the correct value of the first four central moments.
7. For a distribution of 250 heights, calculations showed that the mean, S.D., β_1 and β_2 were 54 inches, 3 inches, 0 and 3 inches respectively. It was however, discovered on checking that the two items 64 and 50 in the original data were wrongly written in place of correct values 62 and 52 inches respectively. Calculate the correct frequency constants.
[**Ans.** $\mu = 54, \sigma^2 = 8.81, \mu_3 = -17.3, \mu_4 = 218.424, \beta_1 = 0.004, \beta_2 = 2.815$]

8. If m and n are positive integers, $m < n$ and if ν_k stands for the absolute moment of a distribution, show that $(\nu_m)^{1/m} \le (\nu_n)^{1/n}$.
9. The variance of symmetrical distribution is 25. What must be the value of μ_4 in order that the distribution be
(a) Leptokurtic, (b) Mesokurtic, (c) Platykurtic [**Ans.** $\mu_4 >, =, < 1875$]

6-40. *Mode or Modal Value of a Distribution*

Let X be a random variable whose density is $f(x)$. The mode of the distribution of X (or *model value* of X) is the variate-value which corresponds to the maximum frequency (probability). [Differential Calculus rules]
If X is continuous, the mode is determined by $x_0, x_1, \ldots$ where $f'(x) = 0$, whose solutions $x_0, x_1, \ldots$ satisfy $f''(x_k) < 0$, $k = 0, 1, 2, \ldots$
If X is discrete, we examine that value of variate which provides maximum probability.

6-41. *Median of a Distribution (Med. X)*

Let a varite X have a c.d.f. $F(x)$. The real number m which satisfies $F(m) \ge \frac{1}{2}$ and $F(m-) \le \frac{1}{2}$ is called a *median* for the distribution of X. Combining these inequalities, median is given by

$$\boxed{\tfrac{1}{2} \le F(m) \le \tfrac{1}{2} + p_m} \qquad [p_m = P(X = m)] \qquad \ldots(1)$$

Obviously, if F is continuous, $p_m = 0$, and median is simply determined by $F(m) = \frac{1}{2}$.
Since there may be many values that satisfy (1), a median is not necessarily unique. Again since (1) always possesses a solution, median always exists.
If F is symmetric c.d.f., the centre of symmetry is clearly the median of the distribution.
Comments. Let X be a *discrete* variate with d.f. 'F'. If $F(x) \ne \frac{1}{2}$, for all admissible values of X, then the *least* permissible value m of X such that $F(m) > \frac{1}{2}$ is taken for median. However, if $F(x_k) = \frac{1}{2}$, then the med X is defined to be $\frac{1}{2}(x_k + x_{k+1})$, where x_k is the least admissible value of X greater than x_k.
Note. In cases where the mean of the distribution does not exist, the importance of median as a centring constant is supreme. (Vide Cauchy distribution).
Example : Let $f(x) = cx$, $c > 0$, $0 \le x \le \sqrt{2/c}$; $f(x) = 0$, elsewhere. If the mode of this distribution is at $x = \sqrt{2}/4$, find the median of this distribution.
Solution. The modal value corresponds to the point where $f(x)$ is maximum. The st. line $y = cx$, is naturally at the greatest height at $x = \sqrt{2/c}$. Thus by hypothesis, $\sqrt{2/c} = \sqrt{2}/4$ which yields $c = 16$. Let m = Med X. Then by definition :

$$\tfrac{1}{2} = P(X \le m) = \int_0^m 16x\,dx = 8m^2 \Rightarrow m^2 = \tfrac{1}{16} \text{ or } m = +\tfrac{1}{4}, (x \nless 0).$$

6-42. Mean (absolute) Deviation (M.a.D.)

The mean absolute deviation about a point $X = a$ may be denoted by $\delta(a)$, and is defined by

$$\delta(a) = \mathsf{E}(|X - a|).$$

We are generally interested in $\delta(\mu)$; or in $\delta(m)$, where m is median. [Notation is non-standard]

Theorem S.D. $\nless$ M.D.

For any distribution, the S.D. σ (say) is not less than mean deviation (M.D.) from the mean.

Proof. Let $\mu = \mathsf{E}(X)$, and $M = \mathsf{E}\,|X - \mu|$, then for any $t \in \mathbb{R}$,

$$\mathsf{E}[\,|X - \mu| + t]^2 \geq 0 \Rightarrow \mathsf{E}(X - \mu)^2 + 2\mathsf{E}\,|X - \mu|\,.\,t + t^2 \geq 0 \;\Rightarrow\; \sigma^2 + 2Mt + t^2 \geq 0.$$

For the non-negative quadratic form in t, the discriminant is negative. Hence

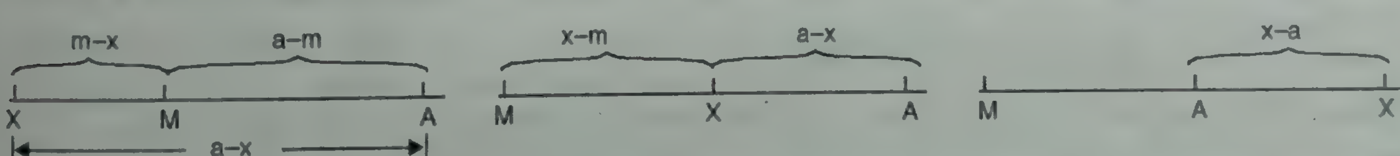
$$M^2 < \sigma^2 \text{ is } \;\sigma \ngtr M.$$

Note. This result trivially follows from Cauchy-Schwartz Inequality. [§10-14]

6-43. Minimal Property of the Median

The sum of the absolute deviations from the median is a minimum.

Proof. Let m = med X. Then we are required to show that $\dot{\mathsf{E}}(\,|X - a|)$ is minimized when $a = m$. That is $\mathsf{E}(\,|X - m|) \leq \mathsf{E}(\,|X - a|)$, or $\mathsf{E}(|X - a| - |X - m|\,) \geq 0$ for any a. In the following, the nature of X is immaterial (discrete or continuous). Firstly, we suppose that $a > m$ and let $d = a - m\ (> 0)$. Define : $A \equiv |X - a| - |X - m|$, and observe that : $-d \leq A \equiv d,\ X \leq m$, [Fig. (a)]; $-d \leq A \leq d,\ m \leq X \leq a$. [Fig. (b)] ; $-d \leq A = -d,\ X \geq a$ [Fig. (c)]



(a) (b) (c)

We observe another random variable W which is related to the three possibilities above. It is given by

$$W = d, \text{ for } X \leq m\ ;\ W = -d, \text{ for } X > m$$

and summarizes the three situations above as $A \geq W$, whence $\mathsf{E}(A) \geq \mathsf{E}(W)$. Now

$$\mathsf{E}(W) = P(X \leq m)\,.\,d + P(X > m)\,.\,(-d) = [P(X \leq m) - P(X > m)]\,d = [2F(m) - 1]\,d$$

since $P(X > m) = 1 - F(m)$. As m = mex X, $F(m) \geq \frac{1}{2}$, and hence $\mathsf{E}(W) \geq 0 \;\Rightarrow\; \mathsf{E}(A) \geq 0$ i.e.

$$\mathsf{E}(|X - a|) \geq \mathsf{E}(|X - m|). \text{ Thus, } \mathsf{E}(|X - a| \nless \mathsf{E}(|X - m|), \text{ for } a > m.$$

Similarly, $\mathsf{E}(|X - a|) \nless \mathsf{E}(|X - m|)$, for $a > m$.

It follows that, $\mathsf{E}(|X - a|)$ is minimized only if $a = m$ = med X.

Another proof. Let X be a continuous r.v. with density $f(x)$, $a < x < b$; where $a = -\infty$ and $b = +\infty$ is also permissible. Let

$$H = \mathsf{E}(|X - m|) = \int_a^b |X - m|\,f(x)\,dx = \int_a^m (m - x)\,dF(x) + \int_m^b (x - m)\,dF(x).$$

Assuming that differentiation under integral sign (DUIS) is valid, we get

$$\frac{\partial H}{\partial m} = \int_a^m f(x)\,dx - \int_m^b f(x)\,dx, \frac{\partial^2 H}{\partial m^2} = f(m) + f(m) = 2f(m) > 0 \quad [\because f \text{ is the density}]$$

$$\frac{\partial H}{\partial m} = 0 \Rightarrow \int_a^m f(x)\,dx = \int_m^b f(x)\,dx = \frac{1}{2} \Rightarrow m = \text{med}\,(X)$$

Since $(\partial^2 H/\partial m^2) > 0$,at $x = m$ (vacuously !), it follows that H is minimized at $x = \text{med}\,(X)$.

Remark. If we make an appeal to Riemann-Stieljes integral, a unified approach combines the cases when X is discrete or continuous. The above proof practically remains unaltered. (See also Ex. 6-11, p.251).

6-50. *Symmetric Density Function and its Applications*

A function f is symmetric about $x = a$, if $f(a + x) = f(a - x)$. A random variable X is symmetric about $X = a$, if its density is symmetric about $X = a$.

Theorem : E(X) = Med (X), if X is Symmetric

If a variate X has symmetric density and if $\mathrm{E}(X)$ exists, then Mean (X) = Median (X).

Proof. Here $\quad f(a + x) = f(a - x), \quad$ [Def. of symmetric density] ...(1)

$$\mathrm{E}(X-a) = \int_{-\infty}^{\infty} (x-a)\,f(x)\,dx = \int_{-\infty}^{\infty} z f(a+z)\,dz = \int_{-\infty}^{\infty} z f(a-z)\,dz \quad [z = (x-a) \text{ and } (1)]$$

$$= \int_{-\infty}^{\infty} (a-t)\,f(t)\,dt = -\int_{-\infty}^{\infty} (t-a)\,f(t)\,dt \quad [t = a - z]$$

$$= -\mathrm{E}(X-a) \quad [\text{Def. of E}]$$

Thus $2\,\mathrm{E}(X - a) = 0 \Rightarrow \mathrm{E}(X) = a$. [The point of symmetry is the mean value]

We now attend to the median. So using $f(a + y) = f(a - y)$, we have

$$P(X > a) = \int_a^{\infty} f(x)\,dx = \int_0^{\infty} f(a+y)\,dy = \int_0^{\infty} f(a-y)\,dy = \int_{-\infty}^{a} f(z)\,dz = P(X < a)$$

$[x = a + y$ and $a - y = z]$. Now $P(X > a) = P(X < a) \Rightarrow$ med $X = a$.

Odd-ordered Central Moments of a Symmetric Distribution

If a Dist. is symmetric about μ, then all its existing central moments of odd order are zero.

Proof. Let f be the p.d.f. of X. By definition of symmetry : $f(\mu - x) = f(\mu + x)$ ...(1)

$$\mu_{2n+1} = \int_{-\infty}^{\infty} f(x)(x-\mu)^{2n+1}\,dx = \int_{-\infty}^{\infty} t^{2n+1} f(\mu+t)\,dt, \quad [x - \mu = t]$$

$$= \int_0^{\infty} t^{2n+1} f(\mu - t)\,dt + \int_0^{\infty} t^{2n+1} f(\mu+t)\,dt \quad [\text{by } (1)]$$

$$= -\int_0^{\infty} \tau^{2n+1} f(\mu+\tau)\,d\tau + \int_0^{\infty} t^{2n+1} f(\mu+t)\,dt, \quad [t = -\tau]$$

We change the dummy τ to dummy t, this results in $\mu_{2n+1} = 0$.

6-51. Worked-out Problems

***Example* 1.** Show that the mean deviation about the median is the minimum.

Solution. Assume that X is continuous and that med $X = m$. If c is any other values, we want to prove that

$$\mathrm{E}(|X-m|) \le \mathrm{E}(|X-c|), \text{ i.e. } H \equiv \mathrm{E}(|X-c| - \mathrm{E}(|X-m|) \ge 0. \qquad ...(1)$$

Firstly, let us assume that $c > m$. Now

$$H = \int_{-\infty}^{\infty} (|x-c| - |x-m| f(x)\,dx = \int_{-\infty}^{\infty} g(x) f(x)\,dx. [g(x) = |x-c| - |x-m|]$$

$$= \left(\int_{-\infty}^{m} + \int_{m}^{c} + \int_{c}^{\infty}\right) g(x)\,dF(x).$$

Now $g(x) = (c-x) - (m-x) = c - m$, when $-\infty < x < m < c$.

$g(x) = (c-x) - (x-m) = (c+m) - 2x$, when $m < x < c$. $[x < c \Rightarrow c + m - 2x > (m-c)]$

$g(x) = (x-c) - (x-m) = m - c$, when $m < c < x < \infty$.

$$\therefore \quad H = \int_{-\infty}^{m} (c-m)\,dF(x) + \int_{m}^{c} (c+m-2x)\,dF(x) + \int_{c}^{\infty} (m-c)\,dF(x)$$

$$\ge \int_{-\infty}^{m} (c-m)\,dF(x) + \int_{m}^{c} (m-c)\,dF(x) + \int_{c}^{\infty} (m-c)\,dF(x) = (c-m)\int_{-\infty}^{m} dF - (c-m)\int_{m}^{\infty} dF(x)$$

$$= (c-m)[P(X \le m) - P(X > m)] = (c-m)[2F(m) - 1] \ge 0. \quad [F(m) \ge \tfrac{1}{2} \text{ as } m = \text{Med } X]$$

From $H \ge 0$, assertion (1) follows. The proof for $m > c$ is exactly similar.

Since c is arbitrary, it follows that $\mathrm{E}(\,|X-m|\,)$ is the minimum.

Furthermore, $\mathrm{E}(\,|X-m|\,) = \mathrm{E}(\,|X-c|\,)$ iff $H = 0$, i.e. iff $m = c$.

Remark. The above proof avoids differentiation under integral sign (DUIS).

***Example* 2.** Show that $\mu_{2n+1} = 0$ does not necessarily imply that the probability distribution is symmetric. Does this result hold for 2-point distribution ?

Solution. (i) Consider three-point distribution : $f(1) = \frac{1}{2}$, $f(-4) = \frac{1}{3}$, $f(5) = \frac{1}{6}$.

$$\mathrm{E}(X) = \frac{1}{2} - \frac{4}{3} + \frac{5}{6} = 0, \quad \mathrm{E}(X^3) = \frac{1}{2} - \frac{64}{3} + \frac{125}{6} = 0.$$

Thus, $\mu_3 = \mathrm{E}(X^3) = 0$, but f has no symmetry about $\mu = 0$.

(ii) Assume that $\mu_3 = 0$ and consider 2-pt distribution $f(a) = p$, $f(b) = q$, $(p + q = 1)$

Here $\mu = ap + bq$. Since $a - \mu = (a-b)q$, $b - \mu = (b-a)\,p$,

$$\therefore 0 = \mu_3 = \mathrm{E}(X-\mu)^3 = (a-\mu)^3 p + (b-\mu)^3 q = (a-b)^3 pq^3 + (b-a)^3 qp^3 = (a-b)^3 pq(q^2 - p^2)$$

$$= (a-b)^3 pq(q-p).$$

Since $a \ne b$, this gives $q - p = 0 \Rightarrow p = q = \frac{1}{2}$. This shows that X is symmetric.

Problems with Solutions Provided at the End of the Text

1*. Find the M.D. from the mean and S.D. of the series ; $a, a+d, a+2d, ..., a+2nd$, and verify that the latter is greater than the former.

2*. Show that M, the M.a.D. about mean m of the variate X, the frequency of whose ith element x_i is f_i, is given by

$$M = \frac{2}{N}\left[m \sum_{x_i < m} f_i - \sum_{x_i < m} f_i\, x_i\right], \qquad \left[N = \sum_i f_i\right]$$

3*. Find the **trimean** (mean, mode and median) for the variate X defined by

$f(x) = \left(\frac{1}{2}\right)^x$, $x = 1, 2, 3, \ldots$ (positive integers).

6-60. *Means and Partition Values*

Let f be the p.d.f. of a given variate X. The following definitions are often needed :

1. ***Arithmetic mean.*** It is denoted by μ, and is defined by

$$\mu = \mathrm{E}(X) = \int_{-\infty}^{\infty} xf(x)\,dx. \qquad [\mathrm{E}(|X|) < \infty]$$

2. ***Geometric mean.*** It is denoted by G, and is defined by

$$\ln G = \mathrm{E}[(\ln X)] = \int_{-\infty}^{\infty} (\ln x)\, f(x)\,dx. \qquad [\mathrm{E}|\ln(X)| < \infty]$$

3. ***Harmonic mean.*** It is denoted by H, and is defined by

$$\frac{1}{H} = \mathrm{E}\left(\frac{1}{X}\right) = \int_{-\infty}^{\infty} \frac{1}{X} f(x)\,dx. \qquad [\mathrm{E}|X^{-1}| < \infty]$$

4. ***Quartiles.*** Suppose the range of the variate X is $[a, b]$. The value of the variate X written Q_r, is called the rth quartile if

$$P\{X \le Q_r\} = \int_a^Q f(x)\,dx = \frac{r}{4}, \quad r = 1, 2, 3.$$

Q_1 and Q_2 are respectively called the first (lower) and the third (upper) quartiles. The quantity $Q_3 - Q_1$ is called the *Quartile deviation* or *Quartile range.* Q_2 is the median (middle quartile).

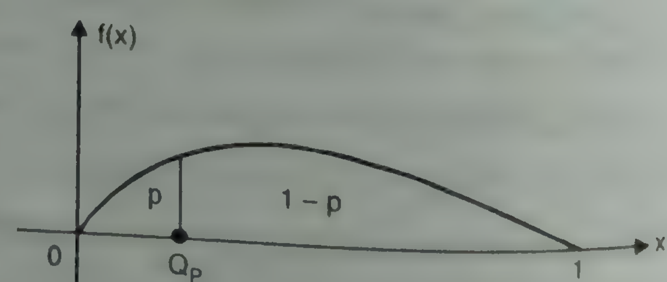
5. ***Deciles.*** Suppose the range of the variate X is $[a, b]$. The value of the variate X, written D_r, is called the rth decile if

$$P\{X \le D_r\} = \int_a^{D_r} f(x)dx = \frac{r}{10}, \quad r = 1, 2, 3, \ldots, 9.$$

6. ***Quantiles of order p.*** The number x satisfying the double inequality

$$p \le F(x) \le p + P_x, \qquad (0 < p < 1).$$

is called the quantile of order p and may be denoted by x_p or Q_p. If $[p_x = P(X = x)]$ $p_x = 0$, e.g. when X is of continuous type, then quantile of order p is a solution of $F(x) = p$. If F is strictly increasing, this Eq. has a unique solution. Otherwise, there may be many solutions of the above Eq. and each one is called *quartile of order p*.



Note. $p = \frac{1}{2}, \frac{1}{4}, \frac{3}{4}, \frac{1}{10}, \ldots,$ provide the values of median, lower quartile, upper quartile, first decile etc.

7. ***Coefficient of variation (C.V.).*** The *relative variance* is σ^2/μ^2 and its positive square root (σ/μ) is called **coefficient of variation.**

6-61. Worked-out Problems

***Example* 1.** Show that in a discrete distribution if the deviations are small compared with the mean M, so that $(x/M)^3$ and higher powers of (x/M) may be neglected, then

(a) $G = M[1-(\sigma^2/2M^2)]$, (b) $M^2 - G^2 = \sigma^2$, (c) $H = M[1-(\sigma^2/M^2)]$,

(d) $H.M = G^2$, (e) $M - 2G + H = 0$, (f) $\text{C.V.} = [2(M-G)/M]^{1/2}$.

Solution. Here $x = X - M$ is the deviation (from Mean) and $(X - M)/M$ is small, $\text{E}(x) = \text{E}(X - M) = 0$.

$$\text{Var}(x) = \text{Var}(X - M) = \text{Var}(X) = \sigma^2 = \text{E}(x^2). \qquad ...(1)$$

We write, $X = (X - M) + M = x + M = M[1 + (x/M)]$ ...(i)

(a) $$\ln G = \text{E}(\ln X) = \text{E}\{\ln M[1 + x/M)]\} = \text{E}\{\ln M + \ln(1 + x/M)\}$$

$$= \ln M + \text{E}[(x/M) - (x^2/2M^2) + ...] = \ln M - (\sigma^2/2M^2) \quad \text{[by lin E and using (1)]}$$

$$\therefore \ln(G/M) = -(\sigma^2/2M^2) \Rightarrow G/M = e^{-\sigma^2/2M^2}$$

or $$G = M[1-(\sigma^2/2M^2)]. \quad \text{[by exponential series]} \qquad ...(2)$$

(b) Squaring both sides of (2), neglecting σ^4/M^4, we get

$$G^2 = M^2(1-\sigma^2/M^2) = M^2 - \sigma^2 \Rightarrow M^2 - G^2 = \sigma^2 \qquad ...(3)$$

(c) $$(1/H) = \text{E}(1/X) = \text{E}\{M^{-1}[1+(x/M)]^{-1}\} = M^{-1}\,\text{E}\{1 - x/M + x^2/M^2 - ...\} \quad \text{[by (i)]}$$

$$= M^{-1}\{1+\sigma^2/M^2\}, \quad \text{[by (1)]}$$

$$H = M[1+(\sigma^2/M^2)]^{-1} = M[1-(\sigma^2/M^2)]. \qquad ...(4)$$

(d) Eliminating σ^2 between (3) and (4), we get $MH = G^2$.

(e) Using the values of G and H from (2) and (4) we get

$$M - 2G + H = M - 2M[1-(\sigma^2/2M^2)] + M[1-(\sigma^2/M^2)] = 0.$$

(f) By definition $C.V. = \sigma/M$. Now from (2)

$$G = M - (\sigma^2/2M) \Rightarrow \sigma^2 = 2M(M-G)$$

$$\therefore \quad \text{C.V.} = \sqrt{2M(M-G)}/M = [2(M-G)/M]^{1/2}.$$

***Example* 2.** In a continuous distribution with $f(x) = y_0 x(2 - x)$, $0 \le x \le 2$ find y_0, mean variance, β_2 H, G, median, M.D. and mode. Show also that $\mu_{2n+1} = 0$.

Solution. The constant y_0 can be determined by normality. Now

$$\text{E}(X^r) = y_0\int_0^2 x^{r+1}(2-x)\,dx = \frac{y_0\, 2^{r+3}}{(r+2)(r+3)}.$$

To find y_0, put $r = 0$ so that $1 = 8y_0/6 \Rightarrow y_0 = \frac{3}{4}$. Thus

$$\mu_r' = \text{E}(x^r) = 2^r.6/(r+2)(r+3), \quad r = 1, 2, 3, ... \qquad ...(1)$$

In particular, $\mathsf{E}(X) = 1$, $\mathsf{E}(X^2) = \frac{6}{5}$, so that Var $(X) = \mathsf{E}(X^2) - \mathsf{E}^2(X) = (\frac{6}{5}) - 1 = (\frac{1}{5})$.

Now $\mu_{2n+1} = \mathsf{E}(X-1)^{2n+1} = y_0 \int_0^2 (x-1)^{2n+1}\, x(2-x)\, dx = y_0 \int_{-1}^{1} t^{2n+1}(1-t^2)\, dt = 0$, $[t = x-1]$

$$\mu_{2n} = \mathsf{E}(X-1)^{2n} = y_0 \int_0^2 (x-1)^{2n}\, x(2-x)\, dx = y_0 \int_{-1}^{1} t^{2n}(1-t^2)\, dt$$

$$= \frac{3}{2}\int_0^1 (t^{2n} - t^{n+2})\, dt = \frac{3}{(2n+1)(2n+3)}.$$

This verifies $\mu_2 = \text{Var}(X) = \frac{1}{5}$. Also $\mu_4 = \frac{3}{35}$, so $\beta_2 = \mu_4/\sigma^4 = \frac{15}{7}$.

From (1) $H^{-1} = \mu'_{-1} = \frac{3}{2} \Rightarrow H = \frac{2}{3}$

$$\ln G = y_0 \int_0^2 (2x - x^2) \ln x\, dx$$

$$\therefore \quad \frac{4}{3}\ln G = \left[\left(x^2 - \frac{x^3}{3}\right)\ln x\right]_0^2 - \int_0^2 \left(x - \frac{x^2}{3}\right) dx = \frac{4}{3}\ln 2 - \frac{10}{9}$$

Thus $$G = 2 \exp(-\tfrac{5}{6}).$$

Let m = med X, then by its definition

$$\frac{1}{2} = \frac{3}{4}\int_0^m (2x - x^2)\, dx \Rightarrow \frac{2}{3} = m^2 - \frac{m^3}{3} \Rightarrow m^3 - 3m^2 + 2 = (m-1)(m^2 - 2m - 2) = 0.$$

The value $m = 1 \in]\,0, 2\,[$, hence med $X = 1$.

M.D. $= \mathsf{E}(|X-1|) = y_0 \int_0^2 |x-1|\, x(2-x)\, dx = y_0 \int_{-1}^{1} |t|(1-t^2)\, dt = \frac{3}{2}\int_0^1 (t - t^3)\, dt = \frac{3}{8}$.

Mode. Differentiating $f(x) = y_0(2x - x^2)$ we get $f'(x) = y_0(2 - 2x)$, $f''(x) = -2y_0$

$\therefore \quad f'(x) = 0 \Rightarrow x = 1 \in]0, 2[$, and $f''(1) < 0$. Hence Mode $X = 1$.

Since $\mathsf{E}(X)$ = Med (X) = Mode $(X) = 1$, the distribution is symmetric about $X = 1$. This is evident from $\mu_{2n+1} = 0$ as well, that yields $\beta_1 = \mu_3^2/\mu_2^3 = 0$. Further

$$(1/H) = \mathsf{E}(X^{-1}) = \tfrac{3}{2} \Rightarrow H = \tfrac{2}{3} \qquad \text{[by (1)]}$$

Example 3. For the triangular (or Simpson) distribution with p.d.f.

$$f(x) = 2(b+x)/b(b+a), \quad -b \le x < 0; \quad f(x) = 2(a-x)/a(a+b), \quad 0 \le x \le a.$$

find tri-mean and variance. Show further that if terms of order $(a-b)^2/a^2$ are neglected, then

$$\text{Mean} - \text{Median} = \tfrac{1}{4}(\text{Mean} - \text{Mode}).$$

Solution. Curve $y = f(x)$ meets y-axis at the point $C(0, 2/(a+b))$ and area $\Delta OBC = b/(a+b)$; area $\Delta OAC = a/(a+b)$. Clearly under $y = f(x)$, $-b \le x \le a$. Area is unity, as expected. Further, mode being the value of X for which $f(x)$ is maximum, is given by $x = 0$. (Fig. *p*.) Now,

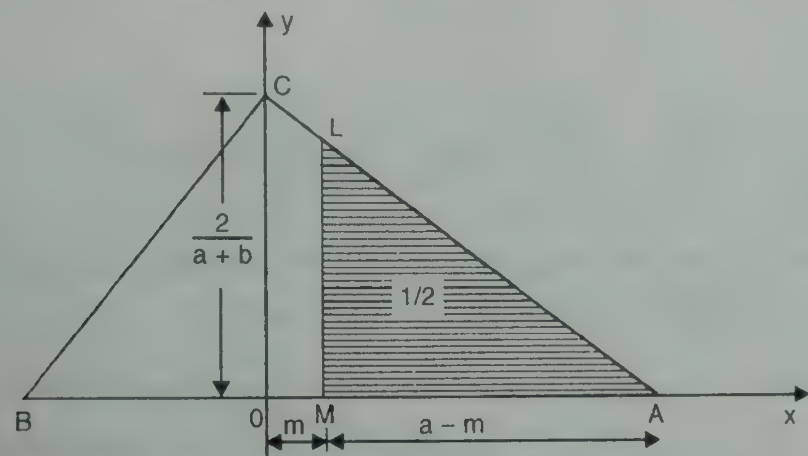
$$E(X^r) = \int_{-b}^{0} \frac{2(b+x)}{b(b+a)} x^r \, dx + \int_0^a \frac{2(a-x)\,x^r}{a(a+b)} dx$$

$$= \frac{2}{b(a+b)}\left[\frac{bx^{r+1}}{r+1} + \frac{x^{r+2}}{r+2}\right]_{-b}^{0} + \frac{2}{a(a+b)}\left[\frac{ax^{r+1}}{r+1} - \frac{x^{r+2}}{r+2}\right]_0^a = \frac{2[a^{r+1} + (-1)^r\, b^{r+1}}{(r+1)\,(r+2)\,(a+b)}$$

Thus, $E(X) = 2(a^2 - b^2)/6(a + b) = (a - b)/3$. $E(X^2) = 2(a^3 + b^3)/12(a + b) = (a^2 - ab + b^2)/6$

$$\text{Var}(X) = (\tfrac{1}{6})(a^2 - ab + b^2) - (\tfrac{1}{9})(a^2 - 2ab + b^2) = (\tfrac{1}{18})(a^2 + ab + b^2).$$

Let m be the median and suppose that $a > b$, so that $m > 0$, $(0 < m < a)$ and hence



$$\frac{1}{2} = \int_m^a \frac{2(a-x)}{a(a+b)} dx = \frac{1}{a\,(a+b)}[2ax - x^2]_m^a = \frac{(a-m)^2}{a(a+b)}.$$

This gives, $a - m = \pm[\tfrac{1}{2}a(a+b)]^{1/2}$ or $m = a \pm [\tfrac{1}{2}a(a+b)]^{1/2}$.

Since $m < a$, we reject positive radical and so $m = a - [\tfrac{1}{2}a + (a+b)]^{1/2}$. ...(1)

Second Part. If $h = (a - b)/a$, then h^2 is negligible. Now using $b = a(1 - h)$

Mean − Median = $\tfrac{1}{3}(a-b) - a + [\tfrac{1}{2}a + (a+b)]^{1/2} = \tfrac{1}{3}ah - a + a(1 - \tfrac{1}{2}h)^{1/2} = ah/12$ ...(2)

$\tfrac{1}{4}$(Mean − Mode) $= \tfrac{1}{4}[\tfrac{1}{3}(a-b) - 0] = ah/12$ ...(3)

From (2) and (3), the result follows.

Note. We can also determine m (median) geometrically from ΔOAC and ΔMAL.

$$\frac{h}{2/(a+b)} = \frac{a-m}{a}, \text{ area} = \tfrac{1}{2} = \tfrac{1}{2}(a-m)/h, \text{ hence } m \text{ as in (1)}.$$

Problems with Solutions Provided at the End of the Text

1*. Let X be a continuous r.v. with unimodal distribution. $[M_0$ = mode X, M_e = med $X]$. Let $\tau_0^2 = E(X - M_0)^2$, $\tau_e^2 = E(X - M_e)^2$, $(M_0 \neq M_e)$. Show that $\tau_0^2 = \tau_e^2$ iff $\mu = (M_0 + M_e)/2$.

2*. Show that the geometric mean G, of the distribution

$$dF = 6(2 - x)(x - 1)\, dx, \quad 1 \le x \le 2$$

is given by $6\ln(16\,G) = 19$.

3*. The mean and S.D. of a variate X are m and σ respectively. If the deviations are small compared with the value of the mean so that $(X/m)^3$ and higher powers of (X/m) are neglected, show that

$$\text{mean}\,(\sqrt{X}) = \sqrt{m}\left(1-\frac{\sigma^2}{8m^2}\right), \quad \text{mean}\left(\frac{1}{\sqrt{X}}\right) = \frac{1}{\sqrt{m}}\left(1+\frac{3\sigma^2}{8m^2}\right).$$

4*. A variate X has p.d.f. : $f(x) = C/(1+x^2)$, $-\infty < x < \infty$.
Determine C, the distribution function and evaluate $P(X \geq 0)$. Find also mode, the mean, variance and quartile range for X.

5*. A frequency distribution is defined by
(i) $f(x) = x^3, 0 \leq x \leq 1$; $f(x) = 3(2-x)^3, 1 \leq x \leq 2$.
(ii) $f(x) = 3x^3, 0 \leq x \leq 1$, $f(x) = (2-x)^3, 1 \leq x \leq 2$.
Find the mean, the S.D. and the M.D. about the mean.

6*. Let $f(x) = ax^2, 0 \leq x < 1$; $f(x) = a(2-x)^2$, $1 \leq x \leq 2$.
Determine initial and central moments, skewness and excess of r.v. X.

7*. The p.d.f. of (shifted) exponential distribution is

$$f(x) = y_0\, e^{-b(x-a)}, \quad a \leq x < \infty. \quad (a, b, y_0 \text{ are constants}).$$

Show that $y_0 = b = 1/\sigma$ and $a = m - \sigma$, where m and σ are respectively the mean and S.D. of the distribution. Show also that $\beta_1 = 4$ and $\beta_2 = 9$.

8*. In a continuous distribution with $f(x) = y_0\,(x - x^2)$, $0 < x \leq 1$, find S.D., G, H, Q_2, M.D. and Modal value. Is the distribution symmetrical ?

9*. For the distribution : $dF(x) = xe^{-x^2/2a^3}\,dx/a^2$, $0 \leq x \leq \infty$,

show that $[(Q_3 - Q_1)/\sigma]$ is independent of the parameter a.

10*. A continuous distribution over range $]-3, 3[$ has p.d.f.

$$f(x) = a(3+x)^2, \ -3 \leq x \leq -1; f(x) = a(6-2x^2), \ -1 \leq x \leq 1; f(x) = a(3-x)^2, \ 1 \leq x \leq 3.$$

Find the constant a and show that all odd-order simple moments vanish. Find also the variance and the M.D. about mean.

11*. Let $F_X(x) = 1 - (c/x)^a$, if $x > c$, $F(x) = 0$, if $x \leq c$ [Pareto distribution].
Find trimean and Variance of X. Evaluate quantile of order $p = 0.75$ when $a = 3$ and $c = 100$.

Exercise 6(c)

1. Find the mean, variance and quartiles for the distributions :

(a) $dF(x) = 2x\,dx$, $0 \leq x \leq 1$. (b) $dF(x) = \frac{1}{2}\sin x\,dx$, $0 \leq x \leq \pi$.

(c) $dF(x) = \sin x\,dx$, $0 \leq x \leq \pi/2$.

[Ans. $\frac{1}{2}, \frac{1}{8}, \frac{1}{2}, \sqrt{\frac{3}{2}}$ (b) $\pi/2, (\pi^2/2) - 2; \pi/3, 2\pi/3$ (c) $1, \pi - 3, \cos^{-1}(\frac{3}{4}), \cos^{-1}(\frac{1}{4})$]

2. Find the first four central moments for the distributions :

(a) $dF(x) = (\frac{1}{2}a)\,dx, \quad -a \le x \le a.$

(b) $dF(x) = dx, \quad 0 \le x \le 1.$

(c) $dF(x) = ce^{-cx}\,dx, \quad 0 \le x < \infty.$

(d) $dF(x) = [2a/\pi\,(a^2 + x^2)]\,dx \quad -a \le x \le a$

(e) $dF(x) = (\pi^{-1} \sin x)\,dx, \quad 0 \le x \le \pi.$

(f) $f(x) = \begin{cases} x, & 0 \le x \le 1 \\ 2 - x, & 1 \le x \le 2 \end{cases}$

[**Ans.** (a) $= \mu = 0 = \mu_3$, $\mu_2 = a^2/3$, $\mu_4 = a^4/5$. (b) 1/2, 1/12, 0, 1/80. (c) $c^{-1}, c^{-2}, 2c^{-3}, 9c^{-4}$, (d) $0, a^2(4-\pi)/\pi, 0, a^3(3\pi - 8)/3\pi$. (e) $\pi - 4\pi^{-1}, 2 - 16\pi^{-2}, -6\pi + 72\pi^{-1} - 128\pi^{-3}$ (f) 1, 1/6, 0, 1/15]

3. A variate X has the p.d.f. $f(x) = k \sin(\pi x/5)$, $0 \le x \le 5$; $f(x) = 0$ elsewhere. Determine the value of k and show that Med X = 5/2 the quartiles are 5/3 and 10/3 and Var $(X) = 25(\pi^2 - 8)/4\pi^2$.

4. Find the Pearson's coefficients β_1 and β_2 and the harmonic mean for the distributions :
(a) $dF(x) = kx^3 e^{-x}\,dx, \quad 0 < x < \infty.$
(b) $dF(x) = kx^2 e^{-x}\,dx, \quad 0 < x < \infty.$
(c) $dF(x) = kx\,e^{-\lambda x}\,dx, \quad x \ge 0, \ \lambda > 0.$ [**Ans.** (a) 1, 4.5, 3 (b) 4/3, 5, 2 (c) 2, 6, λ^{-1}]

5. For the p.d.f. $f(x) = 1 - |1 - x|$, $0 \le x \le 2$, find μ_3.
Hint. $f(x) = x, 0 \le x \le 1 ; f(x) = 2 - x, 1 \le x \le 2.$

6. For the triangular (or Simpson's) density law $f(x) = y_0\{1 - a^{-1}|x - b|\}$, $|x - b| < a$, show that $y_0 = 1/a$, $\mu = b$ and $\sigma^2 = a^2/6$. Graph the curve for $a = b = 1$.

7. For the distribution $f(x) = 6x(1 - x)$, $0 \le x \le 1$, prove that

$$P\{\mu - 2\sigma < X < \mu + 2\sigma\} = 44\sqrt{5}/25, \quad P\{X < 1/2 | 1/3 < X < 2/3\} = 1/2.$$

Find a number b such that $P(X < b) = 2P(X > b)$.

8. Let $f(x) = 4x^3$, $0 \le x \le 1$ and $f(x) = 0$, elsewhere.
Show that the number a such that the probability that X is equally likely to be greater than or less than a is $(1/2)^{1/4}$. Show also the number b such that the probability that X will exceed b is equal to 0.05 is $(0.95)^{1/4}$.

9. Pareto distribution with parameters r and a is given by the p.d.f.

$$f(x) = ra^r/x^{r+1}, \quad x \ge a ; f(x) = 0, x < a, \quad x > 0.$$

Show that it has a finite nth moment iff $n < r$. Show that the variance of the distribution is $[ra^2/(r-1)^2 \cdot (r-2)]$, $r > 2$.

10. Calculate the S.D. and M.D. from the mean if the p.d.f. is
$f(x) = (3 + 2x)/18, \quad 2 \le x \le 4 ; \quad f(x) = 0$ elsewhere. [**Ans.** $\sigma = 0.57$, $M = 0.49$]

11. A variate X has the density : $f(x) = kx/(1 + x)^3, \quad x \ge 0.$
Prove that $k = 2$, mode = 1/2, median = $\sqrt{2} + 1$. Determine the sign of skewness.

12. Verify for the distribution $f(x) = xe^{-x}$, $(0 \le x < \infty)$, that
Mean – Mode = 3 (Mean – Median). [Use : Root of $2(1 + x) = e^x$ is $x = 1.7$]

13. A continuous variate X has the c.d.f.
$F(x) = 0$, if $x < 1$; $F(x) = k(x - 1)^4$, if $1 \le x \le 3$; $F(x) = 1$, if $x > 3$.

Determine k and the p.d.f. Show that the mean = 13/5 and med $X = 1 + (2)^{3/4}$.

14. Determine k so that the following curve represents a frequency function :
$f(x) = 0\, I\,(x \le -1) + k(x+1)\, I\,(-1 < x \le 3) + 4k\, I\,(3 < x \le 4) + 0\, I\,(x > 0)$.
Show that the value of X about which the mean deviation of this distribution is the least is $(2/\sqrt{3} - 1)$.

15. For the density $f(x) = 1/(1+x)^2$, $0 \le x < \infty$, obtain c.d.f. and hence find the median of the distribution. Also investigate if the mean exists. **[Ans.** $F(x) = 1 - (1+x)^{-1}$, $m_e = 1$, $\mu \to \infty$]

16. A variate X has the c.d.f. : $F(x) = k(ax^b - bx^a)$, $0 \le x \le 1$, $a > b > 1$.

Find k, μ'_r, μ, σ^2 and H. If $a = 2b$, prove that m, the median of the distribution, is given by $m^b = (1 - 2^{-1/2})$. **[Ans.** $k = (a-6)^{-1}$, $\mu'_r = ab/(a+r)(b+r)$, $H = (a-1)(b-1)/ab$]

17. The equation of a frequency curve is $y = y_0 (1 + \frac{1}{2} at)^{[(4/a^2)-1]} e^{-(2t/a)}$, $-(2/a) \le t < \infty$.

Find the constant y_0 and the first four central moments.

$\left[\textbf{Ans.}\ \mu_1' = 0, \mu_2 = 1, \mu_3 = a, \mu_4 = 3 + (3a^2/2).\right]$

[**Hint.** Put $(2/a) = b$, use $x + b = z$ to obtain

$$E(X+b)^r = \int_{-b}^{\infty} (x+b)^r f(x)\, dx = \frac{y_0 (e)^{b^2}}{(b)^{b^2-1}} \int_0^{\infty} e^{-bz} . z^{r+b^2-1}\, dz = \frac{y_0 (e)^{b^2}\, \Gamma(r+b)}{(b)^r\, (b)^{2b^2-1}}$$

$$\text{Now } r = 0 \Rightarrow y_0 = \frac{(b)^{2b^2-1}}{\Gamma(b^2) . e^{b^2}}, \quad E(X+b)^r = \frac{\Gamma(r+b^2)}{(b)^r . \Gamma(b^2)}]$$

18. A function F is defined by $F(x) = a, x \le 0$, $F(x) = b + ce^{-x^2/2}$, $x > 0$

where a, b, c are constants. Determine these constants so that F is the distribution function and find p.d.f. and tri-mean of the distribution.

$\left[\textbf{Ans.}\ F(x) = [1 - e^{-x^2/2}]\, I\,(x > 0) + 0\, I\,(x \le 0), \mu = (\pi/2)^{1/2}, m_e = (2\ln 2)^{1/2}\right]$

19. A variate X has the distribution : $P(X \le x) = 1 - e^{-\lambda x^2}$, $\lambda > 0$, $\le x < \infty$.

Find the median of the distribution and show that $2m_0^2 - m^2 = \sigma^2$, $m_0 = m(2/\pi)^{1/2}$ where m, m_0, σ^2 denote respectively mean, mode and variance of X. Determine also the sign of skewness of the distribution.

20. A variate X has the c.d.f. : $F(x) = 2ax/(a^2 + x^2)$.

Show that $E(X) = a(1 - \ln_e 2)$, $\text{Var}(X) = a^2[(\pi - 3) - (1 - \ln_e 2)^2]$.

Deduce that $E(X) \approx 3a/10$ and $\text{Var}(X) \approx a^2/20$.

21. The c.d.f. F of a variate X is : $F(x) = 1 - \exp(-\alpha \tan x)$, $0 \le x \le \pi/2$, $0 < \alpha < 1$.

Prove that mode $= \frac{1}{2}(\pi - \sin^{-1} \alpha)$, median $= \tan^{-1} (\alpha^{-1} \ln 2)$.

Show also that the differential Eq. satisfied by the mean μ is $(d^2\mu/d\alpha^2) + \mu = \alpha^{-1}$.

22. Target T consists of three parts A, B, C. If a shell hits T, then P (shell hits A) = p, P(shell hits B) = P(shell hits C) = p', $[p + 2p' + 1]$. Two or more shells to hit T simultaneously are ruled out and individual hits at A or B or C are collectively independent events. One hit at A, or one hit at B and one hit at C is enough to destroy T. Let X denote the number of shells whose hitting T destroys it, the lesser number of shells fail to destroy T. Calculate the c.d.f. of X, mean, variance and modal value.

$$\left[\textbf{Ans. } f(1) = p, f(2) = 2p'q' \ldots, f(n) = 2(p')^{n-1} q', \ldots \mu = (2/p') - 1 . \sigma^2 = 2p'/q'\right]$$

6-70. *Moments for Bivariate Distributions*

The (r, s)th joint moment of (X, Y) about the point (a, b) is denoted by $\mu'_{r,s}$ and is defined by

$$\mu'_{r,s} = \mathsf{E}[(X-a)^r (Y-b)^s], \; r = 0, 1, 2, \ldots; s = 0, 1, 2, \ldots \qquad \ldots(1)$$

For the particularly simple point (0, 0), it simplifies to $\mu'_{r,s} = \mathsf{E}(X^r Y^s)$.

Notice that $\mu'_{1.0} = \mathsf{E}(X), \mu'_{0,1} = \mathsf{E}(Y), \mu'_{0.2} = \mathsf{E}(X^2), \mu'_{0.2} = \mathsf{E}(Y^2), \mu'_{1.1} = \mathsf{E}(XY)$, etc.

The (r, s)th joint (mixed) moment of (X, Y) about the point $(\mu'_{1.0}\, \mu'_{1.0}) = (\mu_X, \mu_Y)$ is denoted by $\mu_{r,s}$ and is defined by

$$\mu_{r,s} = \mathsf{E}[(X-\mu_X)^r (Y-\mu_Y)^s] = \mathsf{E}(X_0^r Y_0^s), [X_0 = X - \mu_X, Y_0 = Y - \mu_Y], r\text{'s} = 1, 2, \ldots; \qquad \ldots(2)$$

In particular, $\mu_{2,0} = \mathrm{Var}(X)$, $\mu_{0,2} = \mathrm{Var}(Y)$, $\mu_{1,0} \equiv 0 = \mu_{0,1}$.

6-71. *Covariance and its Operational Properties*

The *covariance* between the r.v.s. X, Y is denoted by Cov (X, Y) or σ_{XY} or μ_1, 1 and is defined by

$$\sigma_{XY} = \mathsf{E}[(X-\mu_X)(Y-\mu_Y)] = \mathsf{E}(X_0 Y_0). \quad [X_0 = X - \mu_X, Y_0 = Y - \mu_Y]$$

1. Cov $(X, Y) = \mathsf{E}(XY) - \mathsf{E}(X)\mathsf{E}(Y)$. [Short-setter]
2. Cov (X, X) = Var (X).
3. Cov (X, Y) = Cov (Y, X) [Commutative law]
4. Cov $(X + Y, Z)$ = Cov (X, Z) + Cov (Y, Z).
5. Cov $(aX + b, cY + d) = ac$ Cov (X, Y).
6. Cov $(aX + bY, cX + dY) = ac\sigma_X^2 + bd\sigma_Y^2 + (ad + bc)\sigma_{XY}$.
7. Cov $(aX + bY, cU + dV) = ac\,\sigma_{XU} + ad\,\sigma_{XV} + bc\,\sigma_{YU} + bd\,\sigma_{YV}$.
8. Cov $(X, Y) = 0$, if X and Y are independent ; converse is not true.

Proofs. Here Lin E denotes "Linearity property of E", $\mu_X = \mathsf{E}(X)$, $\mu_Y = \mathsf{E}(Y)$.

1. Cov $(X, Y) = \mathsf{E}[(X-\mu_X)(Y-\mu_Y)] = \mathsf{E}[XY - X\mu_Y - Y\mu_X + \mu_X\mu_Y]$ [Def. of σ_{XY}]

$$= \mathsf{E}(XY) - \mathsf{E}(X)\mu_Y - \mathsf{E}(Y)\mu_X + \mu_X\mu_Y = \mathsf{E}(XY) - \mathsf{E}(X)\mathsf{E}(Y). \quad \text{[by Lin E]}$$

2. Cov $(X, X) = \mathsf{E}(X_0 . X_0) = \mathsf{E}(X_0)^2 = \text{Var}(X)$.

3. Cov $(X, Y) = \mathsf{E}(X_0 . Y_0) = \mathsf{E}(Y_0 . X_0) = \text{Cov}(Y, X)$.

4. Cov $(X + Y, Z) = \mathsf{E}\,\{[(X+Y) - \mathsf{E}(X+Y)][Z - \mathsf{E}(Z)]\}$

$$= \mathsf{E}\{[(X+Y) - (\mu_X + \mu_Y)][Z - \mu_Z]\} \qquad \text{[by Lin E]}$$

$$= \mathsf{E}[(X_0 + Y_0)Z_0] = \mathsf{E}[X_0 Z_0 + Y_0 Z_0] = \mathsf{E}(X_0 Z_0) + \mathsf{E}(Y_0 Z_0), \qquad \text{[Lin E]}$$

5. Cov $(aX + b, cY + d) = \mathsf{E}\{[(aX+b) - \mathsf{E}(aX+b)][cY+d) - \mathsf{E}(cY+d)]\}$ [by Def]

$$= \mathsf{E}\{[(aX+b) - (a\mu_X + b)][cY+d) - (c\mu_Y + d)]\} \qquad \text{[by Lin E]}$$

$$= \mathsf{E}\{aX_0 . cY_0\} = ac\mathsf{E}(X_0 Y_0) = ac\,\text{Cov}(X, Y). \qquad \text{[by Lin E and def.]}$$

Cor. Cov $(aX, cY) = ac\,\text{Cov}(X, Y)$; $\text{Cov}(X, a) = 0$; $\text{Cov}(a, Y) = 0$; $\text{Cov}(a, b) = 0$.

6. Cov $(aX + bY, cX + dY)$

$$= \mathsf{E}\{[(aX+bY) - \mathsf{E}(aX+bY)].[(cX+dY) - \mathsf{E}(cX+dY)\}, \qquad [\text{Def. } \sigma_{XY}]$$

$$= \mathsf{E}\{[(aX+bY) - (a\mu_X + b\mu_Y)][(cX+dY) - (c\mu_X + d\mu_Y)]\} \qquad \text{[by Lin E]}$$

$$= \mathsf{E}[(aX_0 + bY_0)(cX_0 + dY_0) = \mathsf{E}[acX_0^2 + bdY_0^2 + (bc+ad)X_0 Y_0]$$

$$= ac\,\mathsf{E}(X_0^2) + bd\,\mathsf{E}(Y_0^2) + (bc+ad)\mathsf{E}(X_0 Y_0), \qquad \text{[by Lin E]}$$

$$= ac\,\text{Var}(X) + bd\,\text{Var}(Y) + (bc+ad)\,\text{Cov}(X, Y) \qquad \text{[by Defs.]}$$

Cor. $\quad$ Var $(aX + bY) = a^2\text{Var}(X) + b^2\,\text{Var}(Y) + 2ab\,\text{Cov}(X, Y)$

Letting $c = a$, $d = b$, and using Property, 2, the result follows.

7. Cov $(aX + bY, cU + dV) = \mathsf{E}\{[(aX+bY) - \mathsf{E}(aX+bY)][(cU+dV) - \mathsf{E}(cU+dV)]\}$.

$$= \mathsf{E}\{[(aX+bY) - (a\mu_X + b\mu_Y)][(cU+dV) - (c\mu_U + d\mu_V)]\}, \qquad \text{[by Lin E]}$$

$$= \mathsf{E}[(aX_0 + bY_0)(cU_0 + dV_0) = \mathsf{E}[ac\,X_0U_0 + ad\,X_0V_0 + bc\,Y_0U_0 + bd\,Y_0\,V_0]$$

$$= ac\mathsf{E}(X_0 U_0) + ad\mathsf{E}(X_0 V_0) + bc\,\mathsf{E}(Y_0 U_0) + bd\mathsf{E}(Y_0 V_0), \qquad \text{[by Lin E]}$$

$$= ac\sigma_{XU} + ad\sigma_{XV} + bc\sigma_{YU} + bd\sigma_{YV}. \qquad \text{[by Def. of Cov]}$$

Cor. Let $U = X$ and $V = Y$, the result (6) follows. In fact, all properties from (2) to (6) flow out instantly by suitably defining the constants and the variables.

8. Cov $(XY) = \mathsf{E}(X_0 Y_0) = \mathsf{E}(X_0)\mathsf{E}(Y_0) = 0[X\,\&\,Y \text{ indep and } \mathsf{E}(X - \mu_X) = \mathsf{E}(X) - \mu_X = \mu_X - \mu_X = 0]$.

The converse is not true, see Example 5-16.

6-72. *Covariance of Linear Combinations*

Let $X_1, X_2, \ldots, X_m$ and $Y_1, Y_2, \ldots, Y_n$ be jointly distributed random variables with finite variances. Further, let $\sigma_{ij} = \text{Cov}(X_i, Y_j) . \sigma_i^2 = \text{Var}(X_i)$ and define

$S = a_1X_1 + a_2X_2 + \ldots + a_mX_m$, $T = b_1Y_1 + b_2Y_2 + \ldots + b_nY_n$; then

$$\text{Cov}\ (S, T) = \sum_{i=1}^{m}\sum_{j=1}^{n} a_i b_j \sigma_{ij}. \qquad \ldots(1)$$

$$\text{Var}\ (S) = \Sigma a_i^2 \sigma_i^2 + 2\sum_{i=2}^{m}\sum_{j=1}^{i-1} a_i a_j \sigma_{ij}. \qquad \ldots(2)$$

Proof. We have, with obvious ranges for suffixes i and j

$$\begin{aligned} \mathsf{E}(ST) &= \mathsf{E}(\Sigma\Sigma a_i b_j X_i Y_j) = \Sigma\Sigma a_i b_j\ \mathsf{E}(X_i Y_j), \qquad \text{[by Lin E]} \\ \mathsf{E}(S)\,\mathsf{E}(T) &= [\Sigma a_i \mathsf{E}(X_i)][\Sigma b_j \mathsf{E}(Y_j)] = \Sigma\Sigma a_i b_j \mathsf{E}(X_i)\mathsf{E}(Y_j) \\ \therefore \quad \text{Cov}\ (S, T) &= \mathsf{E}(ST) - \mathsf{E}(S)\mathsf{E}(T) = \Sigma\Sigma a_i b_j\,[\mathsf{E}(X_i Y_j) - \mathsf{E}(X_i)\mathsf{E}(Y_j)] \\ &= \Sigma\Sigma a_i b_j\, \text{Cov}\,(X_i, Y_j). \end{aligned}$$

This prove (1). Now take $m = n$; $X_i = Y_i$, $b_i = a_i$ and the result (2) follows.

6-73. *Covariance Matrix and Correlation Matrix*

Let $\sigma_{ij} = \text{Cov}(X_i, Y_j)$, $\sigma_{ii} = \text{Var}(X_i) = \sigma_i^2$, $\rho_{ij} \equiv \text{Corr}(X_i, X_j)$, $\rho_{ii} = 1$.

The covariance matrix $[\sigma_{ij}]$ and correlation matrix $[\rho_{ij}]$ are defined by

$$[\sigma_{ij}] = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \ldots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \ldots & \sigma_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ \sigma_{n1} & \sigma_{n2} & \ldots & \sigma_{nn} \end{bmatrix} \quad [\rho_{ij}] = \begin{bmatrix} \rho_{11} & \rho_{12} & \cdots & \rho_{1n} \\ \rho_{21} & \rho_{22} & \cdots & \rho_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ \rho_{n1} & \rho_{n2} & \cdots & \rho_{nn} \end{bmatrix}$$

Example : An urn contains a white, b black and c green balls ; ($N = a + b + c$). A ball is drawn from this urn. The random variables X, Y, Z are indicators of the events {a white ball}, {a black ball} and {a green ball}. Evaluate the correlation matrix.

Solution. Write $f(x, y, z) = P\{X = x, Y = y, Z = z\}$. Using definition of indicators [e.g. $X = 1$ if white ball drawn, $X = 0$, otherwise, etc.] we have

$$f(0, 0, 0) = 0, \quad f(1, 0, 0) = a/N, \quad f(0, 1, 0) = b/N, \quad f(0, 0, 1) = c/N$$

$f(1, 1, 0) = f(1, 0, 1) = f(0, 1, 1) = f(1, 1, 1) = 0$. $\mu_x = f(1, 0, 0) = a/N$, $\mu_y = b/N$, $\mu_z = c/N$,

$\sigma_X^2 = f(1, 0, 0) \,.\, f(1, 0, 0) = a/N\,[1 - (a/N)] = a\,(b + c)/N^2$. $\sigma_Y^2 = b(a + c)/N^2$. $\sigma_Z^2 = c(a + b)/N^2$.

$$\sigma_{XY} = \Sigma(x-\mu_x)(y-\mu_y)f(x,y,z) = \left(1-\frac{a}{N}\right)\left(0-\frac{b}{N}\right)\frac{a}{N} + \left(0-\frac{a}{N}\right)\left(1-\frac{b}{N}\right)\frac{b}{N} + \left(0-\frac{a}{N}\right)\left(0-\frac{b}{N}\right)\frac{c}{N} = \frac{-ab}{N^2}.$$

Similarly, $\sigma_{YZ} = -bc/N^2$, $\sigma_{ZX} = ac/N^2$.

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X\sigma_Y} = -\frac{\sqrt{ab}}{\sqrt{(a+c)(b+c)}}, \quad \rho_{YZ} = -\frac{\sqrt{bc}}{\sqrt{(b+a)(c+a)}}, \quad \rho_{ZX} = -\frac{\sqrt{ca}}{\sqrt{(c+b)(a+b)}}$$

$$\therefore \quad [\rho_{ij}] = \begin{bmatrix} 1 & -\sqrt{ab}/\sqrt{(a+c)(b+c)} & -\sqrt{ca}/\sqrt{(c+a)(a+b)} \\ \rho_{YX} & 1 & -\sqrt{bc}/\sqrt{(b+a)(c+a)} \\ \rho_{ZX} & \rho_{ZY} & 1 \end{bmatrix}$$

6-74. A Useful Theorem

Suppose $(X_1, X_2, \ldots, X_n) \underset{\text{i.i.d.}}{\smile} (\mu, \sigma^2)$. If $\overline{X} = \Sigma X_i/n$ then

(a) $E(\overline{X}) = \mu$, (b) $\text{Var}(\overline{X}) = \sigma^2/n$ (c) $\text{Cov}(\overline{X}, X_i - \overline{X}) = 0$.

Proof. We use 'Lin E' repeatedly, if need be :

(a) $E(\overline{X}) = E[(X_1 + X_2 + \ldots + X_n)/n] = (1/n)[E(X_1) + E(X_2) + \ldots + E(X_n)]$

$$= (\mu + \mu + \ldots + \mu)/n = n\mu/n = \mu.$$

(b) $\text{Var}(\overline{X}) = \text{Var}[(X_1 + \ldots + X_n)/n] = (1/n^2)\,\text{Var}[X_1 + \ldots + X_n]$

$$= (1/n^2)[\text{Var}(X_1) + \text{Var}(X_2) + \ldots + \text{Var}(X_n)] = (1/n^2)(\sigma^2 + \sigma^2 + \ldots + \sigma^2]$$

$$= n\sigma^2/n^2 = \sigma^2/n$$

(c)
$$\begin{aligned}
\text{Cov}(\overline{X}, X_i - \overline{X}) &= \text{Cov}(\overline{X}, X_i) - \text{Cov}(\overline{X}, \overline{X}) \\
&= \frac{1}{n}\text{Cov}(\Sigma X_j, X_i) - \text{Var}(\overline{X}) \\
&= \frac{1}{n}\text{Cov}(X_i + \Sigma' X_j, X_i) - [\sigma^2/n], \qquad [\Sigma' X_j \text{ contains no } X_i] \\
&= \frac{1}{n}\{\text{Cov}(X_i, X_i) + \text{Cov}(\Sigma' X_j, X_i)\} - [\sigma^2/n] \\
&= \frac{1}{n}[\text{Var}(X_i) + 0] - (\sigma^2/n) = (\sigma^2/n) - (\sigma^2/n) = 0.
\end{aligned}$$

6.75. Worked-out Problems

Example 1. If X_1, X_2 are independent variates with means μ_1, μ_2 and variances σ_1^2, σ_2^2, then

$$\text{Var}(X_1 X_2) = \sigma_1^2\sigma_2^2 + \mu_1^2\sigma_2^2 + \mu_2^2\sigma_1^2.$$

Deduce that : $\text{Var}(X_1 X_2)/E^2(X_1)\,E^2(X_2) = C_1^2 C_2^2 + C_1^2 + C_2^2$

where $C_i = \sigma_i/\mu_i$, is the coefficient of variation of $X_i = i = 1, 2$.

Solution. Since X_1 and X_2 are independent, $E(X_1 X_2) = E(X_1)\,E(X_2) = \mu_1\mu_2$.

$$E(X_1^2 X_2^2) = E(X_1^2)E(X_2^2) = (\sigma_1^2 + \mu_1^2)(\sigma_2^2 + \mu_2^2). \quad [\mu_2 = \mu_2' - \mu_1'^2]$$

$$\text{Var}(X_1 X_2) = E(X_1^2 X_2^2) - [E(X_1 X_2^2)]^2 = (\sigma_1^2 + \mu_1^2)(\sigma_2^2 + \mu_2^2) - (\mu_1\mu_2)^2$$

$$\therefore \quad \text{Var}(X_1 X_2) = \sigma_1^2\sigma_2^2 + \mu_1^2\sigma_2^2 + \mu_2^2\sigma_1^2. \qquad \ldots(1)$$

Dividing both sides of Eqn. (1) by $\mu_1^2\mu_2^2$ we get

$$\text{Var}(X_1 X_2)/E^2(X_1)\,E^2(X_2) = C_1^2 + C_2^2 + C_2^2 + C_i^2.$$

Note. $\text{Var}(X\,Y) = \text{Var}\,X\,\text{Var}\,Y \Leftrightarrow \mu_X = 0 = \mu_Y$. [Ind (X, Y)]

***Example* 2.** A box contains $N = 2^n$ tickets among which ${}^nC_k$ tickets bear the number $k[0 \le k \le n]$. A group of m tickets is drawn. Find the mean and variance of the sum S of their numbers.

Solution. There is $\binom{n}{0} = 1$ ticket bearing numbers; $\binom{n}{1} = n$ tickets bear number 1, $\binom{n}{2} = n(n-1)/2$ tickets bear number 2, and so on. Let X_i be the (random) number on the ticket drawn on the ith occasion ($i = 1, 2, ..., m$). Then $S = X_1 + X_2 + ... + X_m$. Now

$$P\{X_i = k\} = \binom{n}{k}\Big/2^n = \binom{n}{k}\left(\frac{1}{2}\right)^n \quad 0 \le k \le n. \qquad ...(1)$$

$$\therefore \quad \mathrm{E}(X_i) = \sum_{k=0}^{n} k.\binom{n}{k}\left(\frac{1}{2}\right)^n = \frac{n}{2^n}\sum_{k=1}^{n}\binom{n-1}{k-1} = \frac{n}{2^n}(1+1)^{n-1} = \frac{n}{2}. \qquad ...(2)$$

$$\therefore \quad \mathrm{E}(S) = m\,\mathrm{E}(X_i) = mn/2.$$

To find Var (S), we require $\mathrm{E}(X_i^2)$ and $\mathrm{E}(X_i X_j)$, $i \ne j$. Now

$$\mathrm{E}(X_i^2) = \Sigma k^2\binom{n}{k}\left(\frac{1}{2}\right)^n = \frac{1}{2^n}\sum_{k=0}^{n}[k(k-1)+k]\binom{n}{k} = \frac{n(n-1)}{2^n}\sum_{k=2}^{n}\binom{n-2}{k-2} + \frac{n}{2^n}\sum_{k=0}^{n}\binom{n-1}{k-1}$$

$$= \frac{n(n-1)}{2^n}(1+1)^{n-2} + \frac{n}{2^n}(1+1)^{n-1} = \frac{n(n+1)}{4} \qquad ...(3)$$

$$\therefore \quad \text{Var}\,(X_i) = [n(n+1)/4] - (n^2/4) = n/4. \qquad ...(4)$$

To find covariance σ_{ij}, we play a little trick. Draw out **all** the N tickets. Then

$$\text{Var}\,S_N = \Sigma\text{Var}\,(X_i) + 2\,\Sigma\,\Sigma\,\text{Cov}\,(X_i, X_j) = N\sigma^2 + N(N-1)\,\sigma_{ij}.$$

Since S_N = const ; Var $(S_N) = 0$, whence $\sigma_{ij} = -\sigma^2/(N-1)$. Using $\sigma^2 = n/4$, $N = 2^n$, we obtain $\sigma_{ij} = -n/4\,(2^n - 1)$.

$$\therefore \quad \text{Var}\,(S) = \sum_{i=1}^{m}\text{Var}\,(X_i) + 2\sum_{i=1}^{m-1}\sum_{j=i+1}^{m}\text{Cov}\,(X_i, X_j)$$

$$= \frac{mn}{4} - 2\binom{m}{2}\frac{n}{4(2^n-1)} = \frac{mn}{4} - \frac{nm\,(m-1)}{4(2^n-1)} = \frac{mn}{4}\left[1 - \frac{m-1}{2^n-1}\right].$$

***Example* 3.** In a lottery, m tickets are drawn at a time out of tickets numbered 1 to n. Find the expectation and the dispersion (variation) of the sum S of the numbers on the tickets drawn.

Solution. Let X_i denote the number on the ith ticket drawn. Then $S = X_1 + X_2 + ... + X_m$. ...(1)
We now evaluate $\mathrm{E}(X_i)$, Var (X_i) and Cov (X_i, X_j). By definitions :

$$\mathrm{E}(X_i) = \Sigma p_i\,x_i = (1/n)\Sigma X_i = (1/n)[1+2+3+...+n] = \tfrac{1}{2}(n+1).$$

$$\mathrm{E}(X_i^2) = \Sigma(1/n)(1^2+2^2+...+n^2) = (n+1)(2n+1)/6.$$

$$\text{Var}\,(X_i) = \mathrm{E}(X_i^2) - \mathrm{E}^2(X_i) = (n^2-1)/2 \qquad ...(2)$$

$$\mathrm{E}(X_i X_j) = \sum_{i \neq j} \frac{x_i}{n} \cdot \frac{x_j}{n-1} = \frac{1}{n(n-1)} (\Sigma x_i) \Big(\sum_{j \neq i} x_j\Big) = \frac{1}{n(n-1)} \Big(\sum_{i=1}^{n} x_i\Big)\Big[\sum_{j=1}^{n} x_j - x_i\Big]$$

$$n(n-1)\,\mathrm{E}(X_i X_j) = (\Sigma x_i)^2 - (\Sigma x_i^2)^2 = [\tfrac{1}{2} n(n+1)]^2 - [n(n+1)(2n+1)/6]^2$$

$$= n(n+1)(3n^2 - n - 2)/12 = n(n+1)(n-1)(3n+2)/12.$$

$\therefore\ \mathrm{E}(X_i X_j) = (n+1)(3n+2)/12.$

$$\therefore\ \mathrm{Cov}(X_i, X_j) = \mathrm{E}(X_i X_j) - \mathrm{E}(X_i)\mathrm{E}(X_j) = \frac{(n+1)(3n+2)}{12} - \frac{(n+1)^2}{4} = -\frac{(n+1)}{12}. \qquad \ldots(3)$$

Thus $\mathrm{E}(S) = m\mathrm{E}(X_i) = \frac{1}{2} m(n+1)$. [Lin E & by (1)]

Now, $\mathrm{Var}(S) = \sum_{i=1}^{m} \mathrm{Var}(X_i) = 2\Sigma\Sigma\,\mathrm{Cov}(X_i, X_j)$ $[j \neq i]$

$$= \frac{m(n^2-1)}{12} + 2\binom{m}{2}\Big(-\frac{n+1}{12}\Big) = \frac{m(n^2-1)}{12} - \frac{m(m-1)}{12}(n+1) \qquad \text{[By (2) \& (3)]}$$

$$= m(n+1)(n-m)/12.$$

Comments. We can trivially find out σ_{ij}. Take out all the tickets.

$$\mathrm{Var}(S_n) = \Sigma\,\mathrm{Var}(X_i) + 2\Sigma\Sigma\,\mathrm{Cov}(X_i, X_j) = n\sigma^2 + n(n-1)\,\sigma_{ij}.$$

Now S_n = const, hence $\mathrm{Var}(S_n) = 0$, whence $\sigma_{ij} = -\sigma^2/(n-1) = -(n+1)/12$.
However, the method above has its own merit, it shows rich evaluations.

Problems with Solutions Provided at the End of the Text

1*. Is it possible to have data : $\mathrm{Var}(X) = 2 = \mathrm{Var}(Y)$, $\mathrm{Cov}(X, Y) = -3$?

2*. Simplify $Q = \mathrm{E}(XY + Y) - \mathrm{E}(X+1)\mathrm{E}(Y)$, when X and Y are not independent.

3*. If X and Y are indicator variates, show that

$\mathrm{Cov}(X, Y) = 0$ iff X and Y are independent.

4*. A deck of n numbered cards is throughly shuffled and cards are inserted into n numbered cells one by one. If card 'k' falls in the cell 'k' we call it a *match*. Find the mean and variance of the number of such matches.

5*. A bag contains n white and m green balls. If k balls are drawn one at a time wihtout replacement, find the variance of the number of green balls drawn.

6*. *Lottery model.* Suppose a box contains N tickets numbered $v_1, v_2, \ldots, v_N$. A random sample of n tickets is drawn without replacement. Find the variance of the sample mean.

Exercise : A bowl contains N chips numbered from 1 to N consecutively. A random sample of size n is drawn wihtout replacement. Show that the variance of the sample mean is $\sigma^2[(N-n)/n(N-1)]$ where σ^2 is the population variance.

Exercise 6(d)

1. Let X and Y be two variates with finite variances. Show that Cov $(X, Y) < \infty$.
2. Let X and Y be two independent non-degenerate variates. Prove that
$$\text{Var}(XY) = \text{Var}(X)\,\text{Var}(Y) \text{ iff } E(X) = 0, E(Y) = 0.$$
3. If X and Y are i.i.d. continuous r.v.s, show that $X - Y$ has a unique median equal to zero.
4. If X and Y are i.i.d. variates, show that $X - Y$ is a symmetric variate. Give an example to show that the result need not be true if X and Y are not independent.
5. Mark true or false with reasons for your answers :
   (a) Cov $(X, Y) = 0 \Rightarrow X$ and Y are independent.
   (b) If Var $(X) >$ Var (Y), then $X + Y$ and $X - Y$ are dependent.
   (c) If $\sigma_X^2 = \sigma_Y^2$ and if $2X + Y$ and $X - Y$ are independent, then X and Y are dependent.
   (d) If Cov $(aX + bY, bX + aY) \neq ab$ Var $(X + Y)$, then X and Y are independent.
6. Find $\mu'_{r,s}$ for the following jointly distributed variates :
   (a) $f(x, y) = 12xy(1-y),\ 0 \le x,\ y \le 1;\ f(x, y) = 0$, elsewhere.
   (b) $f(x, y) = 4x(1-y),\quad 0 \le x,\ y \le 1; f(x, y) = 0$, elsewhere.
   (c) $f(x, y) = \theta^2 x \exp[-\theta x(1+y)], x > 0, y > 0, \theta > 0; f(x, y) = 0$, elsewhere.
   (d) $f(x, y) = 24xy, x > 0, y > 0, x + y \le 1; f(x, y) = 0$, elsewhere.
   (e) $f(x, y) = 6xy(2 - x - y),\ 0 \le x, y \le 1; f(x, y) = 0$, elsewhere.
   [**Ans.** (a) $12[(r+2)(s+2)(s+3)]^{-1}$, (b) $4[(r+2)(s+1)(s+2)]^{-1}$,
   (c) $s!(r-s)!/\theta^r$, (d) $24B(r+2, s+3)/(s+2)$,
   (e) $6(r+2)^{-1}\{(r+4)(r+3)^{-1}(s+2)^{-1} - (s+3^{-1})\}$].
7. Let X denote the number of occurrences of an event A in n independent trials. Find Var (X) by reasonable assumption about probabilities.
8. A train consisting of N boggies arrives at a railway shunting yard. Out of these boggies, n_1 are bound for D_1. (destination) n_2 are bound for $D_2, \ldots, n_k$ are bound for D_k, $(n_1 + \ldots + n_k = N)$. The loggies are coupled at random along the train regardless of their destinations. If two adjacent boggies have the same destination, they are not uncoupled, otherwise they are uncoupled. Let X be the possible number of uncouplings. Find E(X) and Var (X).
9. A freshman dormitory contains 101 black and 300 white students. If room-mates were assigned strictly at random, without regard to colour, show that the mean and variance of the number of students having white room-mates are 303/4 and 14-2 respectively.
10. Two similar decks of n distinct cards each are put into random order and matched against each other. Prove that the probability of having exactly r matches, is given by
$$\frac{1}{r!}\sum_{k=0}^{n-r}(-1)^k/k!,\ r = 0, 1, 2, \ldots, n.$$
Hence prove that the expected number of matches and its variance are equal and are independent of n.
11. An urn contains p^N white and q^N black balls, the total number of balls being N, $(p + q = 1)$. Balls are drawn on by one (without being returned to the run) until a certain number n of balls is reached. Let $X_i = 1$, ith ball drawn is white ; $X_i = 0$, otherwise.

Show that Cov $(X_i, X_j) = -pq/(N-1), i \neq j$. Deduce that Var $(S_n) = npq\,(N-n)\,(N-1)^{-1}$.

12. A die is thrown $(n+2)$ times. After each throw a "+" is recorded for 4, 5 or 6, and a '–' for 1, 2, 3, the signs forming an ordered sequence. To each, except the first and the last sign, is attached a characterising r.v. which assumes the value 1, if both the neighbouring signs differ from the one between them, and 0, otherwise. If $X_1, X_2, \ldots, X_n$ are the said r.v's. Show that Var $(S_n) = (5n-2)/16$.

13. In a lottery containing n numbers $(1, 2, 3, \ldots, n)$; m numbers are drawn at a time. Let X_k represent the frequency of a specified number k in N drawings. Prove that

$$E(X_k) = Np,\ E(X_k - Np)^2 = Npq.\ E[(X_j - Np)(X_k - Np)] = Np(p' - p) \qquad (j \neq k)$$

where $p = m/n$, $q = 1 - p$, $p' = (m-1)/(n-1)$.

14. A box contains k varieties of objects, the number of objects of each variety being the same. These objects are drawn one at a time and put back before the next drawing. Let n denote the smallest number of drawings which produce objects of all varieties. Find $E(n)$ and Var (n) when k is (i) finite, (ii) infinite.

$\left[\textbf{Ans.}\ \text{(i)}\ k\Sigma(1/r),\ k^2\Sigma(1/r^2) - k(\Sigma\, 1/r)\ 1 \leq r \leq k,\ \text{(ii)}\ k\ln k,\ k^2(1-k^{-1}) - k\ln k.\right]$

A Career is born in Public but talent, like leveled water, in Privacy.

(Marilyn)

Everything has its limits, iron ore cannot be educated into gold.

(Mark Twain)

Appendix : Empirical Models

D

D-1. Empirical Models

A probability model is *empirical* if the probabilities assigned to possible events are secured from observations on the phenomenon under study. In the case of a population, these are the actual values of proportions in the population under study. For repeated trials of experiments, these are taken to be observed relative frequencies.

Suppose $x_1, x_2, ..., x_n$ are the observed values of a sample of size n from some population. Then by analogy with theoretical moments, empirical moments are defined as follows :

***Definition* 1.** The rth moment about origin of an empirical distribution is given by

$$m'_r = (\Sigma x_i^r)/n, \quad i = 1, 2, ..., n.$$

Empirical moments are also called *sample moments*, since these are based on sample values. m'_1, often written $\bar{x}$, is called the *sample mean* and serves to measure where the empirical distribution is centred.

***Definition* 2.** The rth moment about the mean of an empirical distribution is given by

$$m_r = [\Sigma(x_i - \bar{x})^r]/n, \quad i = 1, 2, ..., n.$$

The central moment m_2, often written s^2, is called the *sample variance*, and so s is called the *sample standard deviation* (S.D.).

If the observed values $x_1, x_2, ..., x_n$ have been classified as a frequency table with x_i representing the ith class mark, f_i representing the frequency, (i.e. the number of observations ; repetitions) in the ith interval, and h denoting the number of intervals, then the above definitions are modified to

$$m'_r = \frac{1}{n}\sum_{i=1}^{h} f_i\, x_i^r, \quad m_r = \frac{1}{n}\sum_{i=1}^{h} f_i (x_i - x)^r ; \quad n = \sum_{i=1}^{h} f_i.$$

Remarks. The analogy between empirical distribution and theoretical distribution is remarkably preserved if one permits "E(...)" to stand for Σf_i (...)/n or Σp_i (...), i.e. change E to summation with frequency × observed values. Thus, the geometric means G and harmonic mean H may be defined by

$$\ln G = \mathrm{E}(\ln X) = [\Sigma f_i \ln x_i]/n. \quad 1/H = \mathrm{E}[1/X] = [\Sigma f_i (1/x_i)]/n.$$

All other concepts of empirical distribution can be had through the above device from the corresponding concepts of theoretical distributions.

D-2. Mean and Variance of Combined Sets

Given two sets of observations with sizes n_1, n_2 with means m_1, m_2 and with variances σ_1^2 and σ_2^2; required to find the mean and the variance when the two sets are merged together.

Let the elements of the two sets be denoted by the symbols x_i and y_i with corresponding frequencies f_i and g_j ; then the definition of mean supplies :

$$m_1 = \Sigma f_i x_i / n_1 \ ; \ m_2 = \Sigma g_i y_j / n_2 , \ i = 1, 2, ..., \ j = 1, 2, ... \qquad ...(1)$$

Let m and σ^2 be the mean and variance of the composite set. After the merger, the composite set consists of the elements of the type $f_i x_i + g_j y_j$ whose tota is $n_1 + n_2$, so the definition of mean yields.

$$m = \Sigma(f_i x_i + g_j y_j)/(n_1 + n_2) = (n_1 m_1 + n_2 m_2)/(n_1 + n_2). \ \text{[by (1)]} \qquad ...(2)$$

Variance. For the composite set we have to consider the terms of the type $f_i x_i^2 + g_j y_j^2$.

Using the formula $\mu_2 = \mu_2' - \mu_1'^2$, we get at once

$$\sigma^2 = [\Sigma(f_i x_i^2 + g_j y_j^2)]/(n_1 + n_2)] - m^2. \qquad ...(3)$$

From the individual sets :

$$\sigma_1^2 = [\Sigma f_i x_i^2)/n_1] - m_1^2 ; \quad \sigma_2^2 = [\Sigma(g_j y_j)/n_2] - m_2^2$$

$$\therefore \quad \Sigma f_i x_i^2 = n_1 (m_1^2 + \sigma_1^2) \ ; \quad \Sigma g_j y_j^2 = n_2 (m_2^2 + \sigma_2^2) \qquad ...(4)$$

Making the substitutions from (4) and (2) into (3), we get

$$\sigma^2 = \frac{n_1 (m_1^2 + \sigma_1^2) + n_2 (m_2^2 + \sigma_2^2)}{(n_1 + n_2)} - \left(\frac{m_1 n_1 + m_2 n_2}{n_1 + n_2}\right)^2 \Rightarrow \sigma^2 = \frac{n_1 \sigma_1^2 + n_2 \sigma_2^2}{n_1 + n_2} + \frac{n_1 n_2 (m_1 - m_2)^2}{(n_1 + n_2)^2}$$

Extension. For the composite set which comprises three component sets (n_i, m_i, σ_i^2), $1 \le i \le 3$ we readily obtain

$$\sigma^2 = \frac{n_1 \sigma_1^2 + n_2 \sigma_2^2 + n_3 \sigma_3^2}{n_1 + n_2 + n_3} + \frac{n_1 n_2 (m_1 - m_2)^2 + n_2 n_3 (m_2 - m_3)^2 + n_3 n_1 (m_3 - m_1)^2}{(n_1 + n_2 + n_3)^2}$$

Extension to composite set of k component sets is now obvious.

Example : An analysis of daily wages paid to 586 and 648 workers in two firms A and B, belonging to the same industry, reveals an average daily wage of Rs. 52.5 and Rs. 47.5 with variance of wage distribution Rs. 100 and Rs. 121 respectively.

(a) Which firm pays out the larger amount as daily wages ?

(b) In which firm is greater variability in individual wages ?

(c) What are the measure of (i) Average daily wages ?

(ii) The variability in individual wages of all the workers in the two firms taken together.

Solution. The given data, in terms of usual notation is

$$n_1 = 586, \ n_2 = 648, \ m_1 = 52.5, \ m_2 = 47.5, \ S_1^2 = 100, \ s_2^2 = 121.$$

(a) Daily wages paid to the workers by firm A : $m_1 n_1 = 52.5 \times 586 = 30765$ rupees.
Daily wages paid to the workers by firm B : $m_2 n_2 = 47.5 \times 648 = 30780$ rupees.
Obviously, the firm B pays out larger amount per day.

(b) C_A = Coefficient of variation for firm $A = 100\, s_1/m_2 = 100 \times 10/52.5 = 19.$

C_B = Coefficient of variation for firm $B = 100\, s_2/m_2 = 100 \times 11/47.5 = 23.16.$

Since $C_B > C_A$, it follows that there is greater variability in the individual wages of the firm B.

(c) When the two populations are **taken together** :

(i) $m = \dfrac{m_1 n_1 + m_2 n_2}{n_1 + n_2} = \dfrac{30765 + 30780}{585 + 648} = \dfrac{61545}{1234} = 49.87$ rupees

(ii) $s^2 = \dfrac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2} + \dfrac{n_1 n_2 (m_1 - m_2)^2}{(n_1 + n_2)^2} = \dfrac{58600 + 78408}{1234} + \dfrac{586 \times 648 \times 25}{(1234)^2} = \dfrac{137008}{1234} + \dfrac{293 \times 8100}{617 \times 67}$

$= 111.03 + 6.234 = 117.264$

$\therefore \quad s = \sqrt{117.264} = \sqrt{117.3} = 10.82$ rupees.

D-3. *Sheppard's Corrections*

For a grouped data provided with class intervals, we use

$$m_r' = (\Sigma f_i x_i^r)/N, \qquad m_r = [\Sigma f_i (x_i - \bar{x})^r]/N, \qquad N = \Sigma f_i$$

But the mean and variance, and even higher moments of a frequency distribution are calculated on the assumption that all variate-values in each group interval are *exactly* at its centre, which is very roughly true. The resuting errors in the values of moments are reduced by using the so called *Sheppard's Correction* :

$$(m_1')_c = m_1', (m_2')_c = m_2' - (h^2/12)$$

$$(m_3')_c = m_3' - m_1'(h^2/4), \quad (m_4')_c = m_4' - m_2'(h^2/2) + (7/240)h^4$$

$$(m_2)_c = m_2 - (h^2/12), \quad (m_3)_c = m_3$$

$$(m_4)_c = m_4 - (h^2/2)\, m_2 + (7/240)h^4$$

where h is the length of the class-interval and suffix c indicates the corrected value.

Remark. Sheppard's Corrections often tend to *over-correct* and thus replace old errors by the new errors; consequently these are not much useful.

Example : Show that if the range of six times the S.D. contains atleast 18 class-intervals, Sheppard's correction will make a difference of less than 0.5% in the uncorrected value of S.D.

Solution. Let h be the length of class interval, then $6\sigma > 18h$ (by hypothesis), so that $h < \sigma/3$. Let the suffix c indicate the corrected value ; then

$$(\sigma^2)_c = \sigma^2 - (h^2/12) > \sigma^2 - (\sigma^2/108) \qquad [\because h^2 < \sigma^2/9]$$

or $(\sigma)_c > \sigma\left(1-\frac{1}{108}\right)^{1/2} = \sigma\left(1-\frac{1}{216}\right) \Rightarrow \sigma-\sigma_c < \frac{1}{216}\sigma < \frac{1}{200}\sigma = \left(\frac{1}{2}\%\right)\sigma.$

D-4. *Measure of Skewness* : S_k

(a) Pearson's first coefficient of skewness $(S_k)_1$ = (Mean – Mode)/S.D. = $(\bar{x} - \hat{x})/S$.

(b) Pearson's second coefficient of skewness $(S_k)_2$ = 3(Mean – Median)/S.D. = $3(\bar{x} - \tilde{x})/S$.

(c) Bowley's Quartile Coeff. of $S_k = \dfrac{(Q_3-Q_2)-(Q_2-Q_1)}{(Q_3-Q_2)+(Q_2-Q_1)} = \dfrac{Q_3-2Q_2+Q_1}{Q_3-Q_1}$.

(d) 10–90 percentile Coeff. of $S_k = \dfrac{(P_{90}-P_{50})-(P_{50}-P_{10})}{(P_{90}-P_{50})+(P_{50}-P_{10})} = \dfrac{P_{90}-2P_{50}+P_{10}}{P_{90}-P_{10}}$

(e) Moment Coeff. of skewness $\gamma_2 = \mu_3/\sigma^3$.

Behold ! We read $\bar{x}$ (*x*-bar), $\hat{x}$ (*x*-cap), $\tilde{x}$ (*x*-curl) for mean, mode, median

Note on Coefficient of Dispersion (C.D.)

(i) C.D. = $(a - b)/(a + b)$; $a = \max\{x_1, x_2, ..., x_n\}$, $b = \min\{x_1, x_2, ..., x_n\}$

(ii) C.D. = $\frac{1}{2}(Q_3-Q_1)/\frac{1}{2}(Q_3+Q_1)$ (iii) C.D. = S.D./mean.

(iv) C.D. = M.a.D./Average, from which it is computed.

D-5. *Worked-out Problems*

***Example* 1.** For the following data, with given mean = 67.45 find the missing frequencies, Modal value and Bowley's measure of skewness :

Height (inches) :	60-62	63-65	66-68	69-71	72-74	Total
Frequency :	5	18	a	b	8	100

Solution. $a + b + 31 = 100 \Rightarrow a + b = 69$

Since Interpolation formulas for Mode and Quartiles need the boundaries of the intervals, we recast the data as under, using $y = (x - 67)/3$. [Mid-value = 67]

Intervals	*x (Mid-value)*	*y*	*f*	*fy*	*C (cum-freq.)*
59.5 – 62.5	61	–2	5	–10	5
62.5 – 65.5	64	–1	18	–18	23
65.5 – 68.5	67	0	a	0	a + 23
68.5 – 71.5	70	1	b	b	92, ($a + b = 69$)
71.5 – 74.5	73	2	8	16	100
Totals		0	100	b – 12	

From $y = (x - 67)/3$ we get $\bar{y} = (\bar{x}-67)/3 = (0.45)/3 = 0.15$.

Also, $\bar{y} = \Sigma fy/N \Rightarrow (b-12) = 100 \times (0.15) \Rightarrow b = 27$ and hence $a = 42$.

This gives the missing frequencies.

Modal value $\hat{x}$ is given by Interpolation Formula

$$\hat{x} = L_1 + \frac{(\hat{f}-f_1).h}{(\hat{f}-f_1)+(\hat{f}-f_2)}$$

where $\hat{f}$ = Max. frequency (modal class), f_1 = preceding frequency to $\hat{f}$, f_2 = following frequency to $\hat{f}$,

L_1 = Lower boundary of the modal class, h = Width of the modal class.

Here $\hat{f} = a = 42, f_1 = 18, f_2 = b = 27, L_1 = 65.5, \;\; h = 3$

$$\hat{x} = 65.5 + \frac{3(42-18)}{84-45} = 65.5 + 1.85 = 67.35$$

The Quartiles Q_1, Q_2, Q_3 are given by Interpolation Formula

$$Q_i = L_1 + \frac{i.(N/4) - C}{f}.h, \qquad i = 1, 2, 3 \ldots$$

where, L_1 = Lower boundary of the Q_i class, N = total frequency.

C = Less than cum-frequency of the class preceding the Q_i-class.

f = Frequency of the Q_i-class, h = width of the Q_i-class.

Median Q_2. Here $(N/2) = 50$. It falls in Cum-frequency $a + 23 = 65$ class.

$$L_1 = 65.5, \;\; C = 23, h = 3, f = 42.$$

$$Q_2 = 65.5 + [(50 - 23) \times 3/42] = 65.5 + 1.93 = 67.43$$

Upper quartile Q_3. Here $3N/4 = 75$. It falls in Cum-frequency 92 class

$$L_1 = 68.5, h = 3, C = 65, f = 27$$

$$Q_3 = 68.5 + [(75 - 65) \times 3/27] = 68.5 + 1.11 = 69.61.$$

Lower quartile Q_1. Here $N/4 = 25$. It falls in Cum-frequency 65-class

$$L_1 = 65.5, h = 3, C = 23, f = 42.$$

$$Q_1 = 65.5 + [3(25 - 23)/42] = 65.5 + 0.143 = 65.64$$

Bowley's coefficient of skewness S_k is

$$S_k = \frac{(Q_3 - Q_2) - (Q_2 - Q_1)}{(Q_3 - Q_2) + (Q_2 - Q_1)} = \frac{Q_1 + Q_3 - 2Q_2}{Q_3 - Q_1} = \frac{135.25 - 134.86}{69.61 - 65.64} = \frac{0.39}{3.97} = 0.098.$$

Example **2.** In a frequency table, the upper boundary of each class interval has a constant ratio to the lower boundary. Show that the geometric mean G is expressible by

$$\ln G = x_0 + (c/N)\,\Sigma f_i (i-1)$$

where x_0 is the logarithm of the mid-value of the first interval and c is the logarithm of the ratio between the upper and lower boundaries.

Solution. Here $N = \sum f_i$ is the total frequency. $f_i/N = p_i$ gives $\sum p_i = 1$. The frequency table is

$$\text{Variable value} : \;(a, ar)\,(ar, ar^2) \ldots (ar^{i-1}, ar^i) \ldots \;\;(ar^{n-1}, ar^n)$$

$$\text{Frequency} : \;\; f_1, \;\; f_2, \;\; \ldots \;\; f_i, \;\; \ldots \;\; f_n.$$

We are given that $x_0 = \ln(a + ar)/2, c = \ln r$. Now by definition

$$\ln G = \mathrm{E}(\ln X) = \sum_i p_i \ln\left(\frac{ar^i + ar^{i-1}}{2}\right) = \Sigma p_i \ln r^{i-1}\left[\frac{(1+r)a}{2}\right]$$

$$= \Sigma p_i [(i-1)\ln r + \ln[(a+ar)/2]$$

$$= \Sigma p_i [(i-1)c + x_0] = c\Sigma p_i (i-1) + x_0 (\Sigma p_i) = x_0 + (c/N)\Sigma f_i (i-1).$$

***Example* 3.** For the following frequency distribution, find the moments measure of skewness and kurtosis. Is the distribution Leptokurtic ?

Class Limits	*Frequency*
110.0 – 114.9	5
115.0 – 119.9	9
120.0 – 124.9	17
125.0 – 129.9	42
130.0 – 134.9	10
135.0 – 139.9	11
140.0 – 144.9	6

Solution. The class mark x_i is the mid-values of the class limits. We next use codification and use empirical formulas corresponding to theoretical (population) formulas*. Let

$y_i = (x_i - 127.45)/5 \quad \Rightarrow \quad x_i = 127.45 + 5y_i$ [Mid class-value = 127.45]. Now

x_i	y_i	f_i	$f_i y_i$	$f_i y_i^2$	$f_i y_i^3$	$f_i y_i^4$
112.45	–3	5	–15	45	–135	405
117.45	–2	9	–18	36	–72	144
122.45	–1	17	–17	17	–17	017
127.45	0	42	0	00	0	00
132.45	1	10	10	10	10	10
137.45	2	11	22	44	88	176
142.45	3	6	18	54	162	486
Total	0	$N = 100$	0	206	36	1238

Calculations of Sample Moments

$$m_1' = \bar{y} = \frac{\Sigma f y_i}{N} = 0, \qquad m_2' = \frac{\Sigma f_i y_i^2}{N} = \frac{206}{100} = 2.06, \qquad m_3' = \frac{\Sigma f_i y_i^3}{N} = \frac{36}{100} = 0.36.$$

$$m_4' = \frac{\Sigma f_i y_i^4}{N} = \frac{1238}{100} = 12.38$$

Since $m_1' = 0$, Simple Sample moments = Central sample moments for Y.

$* \; X = 127.45 + 5Y \quad \Rightarrow \quad \mathrm{E}(X) = 127.45 + 5\,(\mathrm{E}Y)$

$\mathrm{E}(X - \mu_X)^r = (5)^r\, \mathrm{E}(Y - \mu_Y)^r \Rightarrow (\mu_r)_X = 5^r . (\mu_Y)^r.$

Corresponding for samples : $M_r = (5)^r\, m_r$.

Also $\qquad \alpha_r = \dfrac{\mu_r}{\sigma^r}$ provides $a_r = \dfrac{M_r}{(\sqrt{M_2})^r}$.

For central sample moments $M_r = (5)^r m_r$, hence.

$$M_2 = 25\, m_2 = 25\,(2.06) = 51.5, \quad \sqrt{M_2} = 7.17635$$

$$M_3 = 125\, m_3 = 125(0.36) = 45 \qquad M_4 = 625\, m_4 = 625(12.38) = 7737.5$$

$$\text{Coeff of Skewness, } a_3 = M_3/(\sqrt{M_2})^3 = \frac{45}{(7.176)^3} = 0.122$$

$$\text{Coeff of Kurtosis, } a_4 = \frac{M_4}{(\sqrt{M_2})^4} = \frac{7737.5}{(51.5)^2} = \frac{7737.5}{2652.25} = 2.92$$

Since $a_4 < 3$, the distribution is Platykurtic not Leptokurtic.

Comments. If only α_r is needed and individual moments μ_2, μ_3, μ_4 for X are not required, we shorten evaluations by using

$$\alpha_r = \left(\frac{\mu_r}{\sigma^r}\right)_X = \left(\frac{\mu_r}{\sigma^r}\right)_Y, \quad Y = \frac{X-a}{h}.$$

***Example* 4.** If Pearson $S_k = 0.32$, $\mu = 29.6$, $\sigma = 6.5$ find M_0 and M_e. However, if $M_0 = 24.8$, what will be S.D. ?

Solution.

$$S_k = \frac{\mu - M_0}{\text{S.D.}} \Rightarrow 0.32 = \frac{29.6 - M_0}{6.5} \Rightarrow M_0 = 27.52 \text{ (Mode)}$$

$$S_k = \frac{3(\mu - M_e)}{\text{S.D.}} \Rightarrow 0.32 = \frac{3(29.6 - M_e)}{6.5} \Rightarrow M_e = 28.91. \text{ (Median)}$$

$$S_k = \frac{\mu - M_0}{\text{S.D.}} \Rightarrow 0.32 = \frac{29.6 - 24.8}{\sigma} \Rightarrow \sigma = 15.$$

Problems with Solutions Provided at the End of the Text

1*. Show that the weighted arithmetic mean of the first n natural numbers, whose weights are equal to the corresponding numbers, is equal to $(2n + 1)/3$.

2*. Find the mean and variance of the distribution in which the values of the variate X are positive integers $1, 2, \ldots, n$, the frequency of each being unity.

3*. Show that, if the variate X takes values $0, 1, 2, \ldots, n$ with frequencies proportional to ${}^nC_0, {}^nC_1, {}^nC_2, \ldots, {}^nC_n$ respectively, then the mean of the distribution is $n/2$, mean square deviation about $x = 0$ is $n(n + 1)/4$ and the variance is $n/4$.

4*. From a sample of n observations, the arithmetic mean and variance are calculated. It is then found that one of the values x_1 is in error and that it should be replaced by y_1. Show that the adjustment to the variance to correct this error is

$$\frac{1}{n}(y_1 - x_1)\left[(x_1 + y_1) - \frac{y_1 - x_1 + 2T}{n}\right]$$

where T is the total of the original values.

5*. Show that for a set of positive values of the variate X, the arithmetic, geometric and harmonic means are special cases of the pth root of the mean of the pth power of the variables.

6*. For a set of observations : $x_1, x_2, \ldots, x_n$, let r be the range and let $S, \hat{S}$ be the standard deviations defined by $S^2 = \Sigma(x_i - \bar{x})^2/n$, $\hat{S}^2 = \Sigma(x_i - \bar{x})^2/(n-1)$.
Prove : $\hat{S} \le r\sqrt{n/(n-1)}$; $\quad S \le r$.

Exercises

1. A variate takes the value $a, ar, ar^2, \ldots, ar^{n-1}$ each with frequency unity. If A, G, H are respectively the arithmetic, geometric and harmonic means, show that

$$A = a(1-r^n)/n(1-r), \qquad G = ar^{(n-1)/2}, \quad H = an(1-r)r^{n-1}/(1-r^n).$$

Prove that $G^2 = AH$, and also $A > G > H$ unless $n = 1$, when $A = G = H$.

2. If G_x is the geometric mean of N x's and G_y is the geometric means of N y's, then the geometric mean G of the $2N$ values is given by $G^2 = G_x \cdot G_y$.

3. If $\bar{x}_w$ is the weighted mean of x_i's with weight w_i's, prove that

$$\left(\sum_{i=1}^{n} w_i\right)\left[\sum_{i=1}^{n} w_i\,(x_i - \bar{x}_w)^2\right] = \sum_{i=1}^{n}\sum_{j=1}^{n} w_i\, w_j\,(x_i - x_j)^2 \cdot \left(\sum_{i=1}^{n} w_i \neq 0\right)$$

4. Show that in finding the arithmetic mean of a set of readings on thermometer, it does not matter whether we measure temperature in Centigrade or Fahrenheit, but that in finding the geometric mean it does matter which scale we use.

5. A distribution $x_1, x_2, \ldots, x_n$ with frequencies $f_1, f_2, \ldots, f_n$ is transformed into the distribution $X_1, X_2, \ldots, X_n$ with the same corresponding frequencies by the relation $X_r = ax_r + b$, where a and b are constants. Show that the mean, median and mode of the new distribution are given in terms of those of the first distribution by the same transformation.

6. A set of numbers are all either 0 or 1. If m of the numbers are 1 and n of the numbers are 0, prove that the variance of the set of numbers is $mn/(m+n)^2$.

7. A sample of 35 has mean 80 and S.D. 4. A second sample of 65 has the mean 70 and S.D. 3. Find the S.D. of the combined sample of 100.

8. The first of two samples has 100 items with mean 15 and S.D. 3. If the whole group has 250 items with mean 15.6 and S.D. $\sqrt{13.44}$ show that the mean and S.D. of the 2nd group are 16 and 4 respectively.

9. In a series of measurements, we obtain m_1 values of magnitude x_1, m_2 values of magnitude x_2, and so on. If x is the mean value of all the measurements, prove that the S.D. is

$$[m_r\,(k - x_r)^2 / \Sigma\, m_r]^{1/2} - \delta^2$$

where $\delta = x - k$ and k is any constant.

10. A student obtained the mean and S.D. of 100 observations as 40 and 5.1 respectively. It was later on discovered that he head copied down an observation as 50 instead of 40. Calculate the correct mean and S.D.

11. The arithmetic mean and variance of a set of 10 figures are 17 and 33 respectively. Of the 10 figures, one figure, viz. 26 was subsequently found wrong and was weeded out. What is the resulting mean and S.D. ?

12. The deviation of a distribution is measured from a value differing from the mean of the distribution by x. Show that if x is plotted against the corresponding mean square deviation, the points lie on a parabola.

13. The sum and sum of squares corresponding to length X (in cms) and weight Y (in gms) of 50 tapioca tubers are :

$\Sigma x = 212, \qquad \Sigma x^2 = 902.8, \qquad \Sigma y = 261, \qquad \Sigma y^2 = 1457.6.$

Which is more varying, the length or weight ?

14. Two workers on the same job show the following results over a period of time :

	Worker A	*Worker B*
Mean time to complete work	30	25
S.D. of time for completion	6	4

(a) Which worker appears to be more consistent ?

(b) Which worker appears to be faster in completing the job ?

15. For a group of 200 candidates, the mean and S.D. were found to be 40 and 15 respectively. Later on it was learnt that the scores 43 and 35 were misread as 34 and 53 respectively. Find the corrected mean and S.D. corresponding to the corrected figures.

16. A distribution consists of 3 components with frequencies 200, 250 and 300 having means 25, 10, 15 and S.Ds. 3, 4, 5 respectively. Show that the mean and S.D. of the combined groups are 16 and 7.2 respectively.

Money Can't buy happiness. But it sure makes misery easier to live with.

****

A witty saying proves nothing. (Voltaire)

Conditional Expectation 7

7-10. Conditional Expectation

***Definition* 1.** Let g be a real-valued function of *one variable* x. The conditional expectation of a function $g(X)$, of random variable X, given $A = a < X \le b$, is denoted by $\mathsf{E}[g(X) \mid A]$ and is defined by

$$\mathsf{E}[g(X) \mid a < X \le b] = \int_{-\infty}^{\infty} g(x)\, f_X(x \mid a < x \le b)\, dx \qquad (X : \text{continuous})$$

$$= \int_a^b g(x)\left[f(x) \Big/ \int_a^b f(x)\, dx\right] dx = \int_a^b g(x) \,.\, f(x) dx \Big/ \int_a^b f(x)\, dx \quad \ldots(1)$$

If X is a discrete r.v. we define

$$\mathsf{E}[g(X) \mid a < X \le b] = \sum_{x_i} g(x_i)\, p_X(x_i \mid a < x_i \le b) = \sum_{x_i} g(x_i)\, p(x_i) \Big/ \sum_{x_i} p(x_i), \quad a < x_i \le b. \quad \ldots(2)$$

***Definition* 2.** Let g be a real-valued function of two variables x and y. The conditional expectation of a function g of random variables X and Y, given $Y = y$ is denoted by $\mathsf{E}[g(X, Y) \mid Y = y]$ and is defined by

$$\mathsf{E}[g(X, Y) \mid Y = y] = \begin{cases} \int_{-\infty}^{\infty} g(x, y)\, f_{X \mid Y}(x, y)\, dx, & (X, Y \text{ continuous}) \\ \sum_{x_i} g(x_i, y)\, P(X = x_i \mid Y = y), & (X, Y : \text{discrete}) \end{cases}$$

The cases $g(X, Y) \equiv X$ or $g(X, Y) \equiv Y$ merit special considerations.

***Definition* 3.** Let the two-dim variate (X, Y) have the joint (discrete or continuous) density $f(x, y)$. Let $f_1(x \mid y)$ denote conditional density of X and Y when $Y = y$, and $f_2(y \mid x)$ denote the conditional density of X and Y when $X = x$. The conditional expectation of Y given $X = x$ is denoted by $\mathsf{E}(Y \mid x)$, $[\equiv \mathsf{E}(Y \mid X = x)]$ and the conditional expectation of X given $Y = y$ is denoted by $\mathsf{E}(X \mid y)$, $[\equiv \mathsf{E}(X \mid Y = y)]$. These quantities are defined by

$$\mathsf{E}(Y \mid X = x) = \begin{cases} \int_{-\infty}^{\infty} y f_2(y \mid x)\, dy = \int_{-\infty}^{\infty} \dfrac{y \,.\, f(x, y)}{f_X(x)}\, dy & (X, Y \text{ continuous}) \\ \sum_{y} y f_2(y \mid x) = \dfrac{\Sigma_y\, y f(x, y)}{f_X(x)} & (X, Y : \text{discrete}) \end{cases}$$

$$\mathsf{E}(X \mid Y = y) = \begin{cases} \int_{-\infty}^{\infty} x f_1(x \mid y)\, dx = \int_{-\infty}^{\infty} \dfrac{x\, f(x, y)}{f_Y(y)}\, dy & (X, Y \text{ continuous}) \\ \sum_{x} x f_1(x \mid y) = \dfrac{\Sigma_x\, x f(x, y)}{f_Y(y)} & (X, Y : \text{discrete}) \end{cases}$$

Conditional expectation over intervals are defined as in Definition 1, replacing density by its value over the interval.

***Note* 1.** If X and Y are independent, then $E(X \mid Y = y) = E(X)$, since $f_1(x \mid y) = f_X(x)$

***Note* 2.** $E(Y \mid x)$ is, in general, a function of x and is not a random variable, since x is a particular value of X. However, it can be randomized by letting x to vary.

***Regression curves* :**

(1) The graph of $E(X \mid Y = y) = \phi(y)$, is called the *regression curve* of X on Y.

(2) The graph of $E(Y \mid X = x) = \psi(x)$, is called the *regression curve* of Y on X.

(3) The graph of $E(X \mid y) = \alpha + \beta y$, is called *linear regression* of X on Y.

(4) The garph of $E(Y \mid x) = a + bx$, is called *linear regression* of Y on X.

7-11. *Some Properties of Conditional Expectation*

(i) $E(k \mid X) = k$. (k : constant)

(ii) $E\{(aX + bY) \mid Z\} = aE(X \mid Z) + bE(Y \mid Z)$, [Linear property]

(iii) $E(X \mid Y) = E(X)$, if X is independent of Y.

(iv) $E\{Yg(X) \mid X\} = g(X)\, E(Y \mid X)$, [Pull-through property]

(v) $E\{Y \mid X\,;\, g(X)\} = E(Y \mid X)$.

(vi) $E\{E(X \mid Y\,;\, Z) \mid Y\} = E(X \mid Y)$, [Tower property for multiple-conditioning]

(vii) $\text{Cov}\,(X, \mu_{Y \mid X}) = \text{Cov}\,(X, Y)$.

Linear Property of Conditional Expectation

$$E[ag(X, Y) + b\, h(X, Y) \mid y_0] = aE[g(X, Y) \mid y_0] + bE[h(X, Y) \mid y_0] \quad \ldots(1)$$

where a, b are constants and g, h are real-valued functions of x and y.

Proof. The following is an equivalent statement to result (1).

$$\int_{-\infty}^{\infty} [ag(x, y) + bh(x, y)]\, f_1(x \mid y_0)\, dx = a\int_{-\infty}^{\infty} g(x, y)\, f_1(x \mid y_0)\, dx + b\int_{-\infty}^{\infty} h(x, y)\, f_1(x \mid y_0)\, dy.$$

When X, Y are both discrete, the result follows on replacing integrals by Summations

7-20. *Expectation of a Variate by Conditioning*

Double-Expectation Rule. (Double-E Rule) or Law of Iterated Expectation

$$E[g(X)] = E\{E[g(X) \mid Y]\} \qquad \text{provided } Eg(X) < \infty.$$

Proof. (i) Suppose X and Y are discrete variates, then by definition

$$E\{E[g(X) \mid Y]\} = E\{\Sigma_x\, g(x)\, P(X = x \mid Y = y)\} = \Sigma_y \left[\Sigma_x\, g(x) \cdot \frac{P(X = x, Y = y)}{P_Y(y)}\right] \cdot P_Y(y)$$

$$= \Sigma_x\, g(x) \cdot \Sigma_y P(X = x, Y = y) = \Sigma_x\, g(x) \cdot P(X = x) = E[g(X)].$$

(ii) Suppose X and Y are continuous variates, then by definition

$$E\{E[g(X) \mid Y]\} = E\left\{\int_{-\infty}^{\infty} g(x) f_1(x \mid y)\, dx\right\} = \int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} g(x) \frac{f(x, y)\, dx}{f_Y(y)}\right] f_Y(y)\, dy$$

$$= \int_{-\infty}^{\infty} g(x) \left\{\int_{-\infty}^{\infty} f(x, y)\, dy\right\} dx = \int_{-\infty}^{\infty} g(x) \cdot f_X(x)\, dx = E[g(X)].$$

Special Case :

$$\boxed{E(X) = E[E(X \mid Y)].}$$

Remarks. If $g(X) = \infty$, Double = E Rule is not true. To see this, consider p.d.f. of a variate Y given by

$$f_Y(y) = \frac{(y^{-1/2}\, e^{-y/2})}{\left(\Gamma\left(\frac{1}{2}\right).\sqrt{2}\right)}, \quad y>0; \quad f_y(y) = 0,\ y \le 0.$$

Suppose that the conditional p.d.f. of X given y (> 0) is

$$f_1(x \mid y) = (2\pi)^{-1/2} y^{1/2} e^{-yx^2/2}, \quad -\infty < x < \infty,\ y > 0$$

Since $f(-x \mid y) = f(x \mid y)$, $E(X \mid y) = 0$; so E $E(X \mid y)$ exists, its value is zero.

The marginal p.d.f. of X is given by

$$f_X(x) = \int_0^{\infty} f_Y(y)\, f_1(x \mid y)\, dy = \frac{1}{2\sqrt{\pi}\,\Gamma\left(\frac{1}{2}\right)} \int_0^{\infty} e^{-y(1+x^2)/2} dy = \frac{1}{\pi(1+x^2)} \quad -\infty < x < \infty.$$

This is the Cauchy density for which $E \mid X \mid = \infty$. [Example 1(*e*)].

Thus, $E[E(X \mid Y)]$ exists but $E(X)$ does not exist.

7-21. *Wald's Theorem on Mean and Variance of a Random Sum*

If $X_1, X_2, \ldots, X_N$ are independent variates with identical mean μ and variance σ^2, and if N is a random variable, independent of X_i then setting $S_N = X_1 + X_2 + \ldots + X_N$, we have

$$E(S_N) = E(N)\, E(X) \qquad \ldots(1)$$

$$\text{Var}\,(S_N) = \text{Var}(X)\, E(N) + \text{Var}(N)\, E^2(X). \qquad \ldots(2)$$

Proof. $E(S_N) = E[E(S_N \mid N)]$ [By Double-E Rule] ...(i)

$$E(S_N \mid N = n) = E(S_n) = E(X_1 + X_2 + \ldots + X_n) = n\mu, \quad [E(X_i) = \mu]$$

$$E(S_N) = E(N\mu) = \mu E(N) = E(X)\, E(N) \qquad \ldots(1)$$

Now, $\text{Var}(S_N) = E(S_N^2) - E^2(S_N)$ (Computational formula) ...(3)

$$E(S_N^2) = E[E(S_N^2 \mid N)] \quad \text{[By Double-E Rule]} \qquad \ldots(\text{ii})$$

$$E(S_N^2 \mid N = n) = E(S_n^2) = \text{Var}(S_n) + E^2(S_n) = n\sigma^2 + (n\mu)^2 \quad [\because \mu'_2 = \mu_2 + \mu'^2_1]$$

$$\therefore \quad E(S_N^2 \mid N) = N\sigma^2 + N^2\mu^2 \Rightarrow E(S_N^2) = E(N\sigma^2 + N^2\mu^2) = \sigma^2 E(N) + \mu^2 E(N^2) \qquad \ldots(4)$$

From (1), (3) and (4) we obtain

$$\text{Var}\,(S_N) = \sigma^2 E(N) + \mu^2[E(N^2) - E^2(N)] = \text{Var}(X)\, E(N) + \text{Var}(N) . E^2(X). \qquad \ldots(2)$$

Comments. Results (1) and (2) can be quickly had by MGF method.

7-30. *Formula for Conditional Variance*

$\text{Var}(X) = E\{\text{Var}(X \mid Y) + \text{Var}\{E(X \mid Y)$ = mean (cond. variance) + Var(cond. mean)...(1) ...(A)

Proof. Recall : $\text{Var}(Z) = E(Z^2) - \{E(Z)\}^2$

$$\therefore \quad \text{Var}\,(X \mid Y) = E(X^2 \mid Y) - E^2(X \mid Y) \qquad [\text{Replace } Z \text{ by } X \mid Y]$$

So $E\{Var(X \mid Y)\} = E\{E(X^2 \mid Y)\} - E\{E^2(X \mid Y)\} = E(X^2) - E\{E^2(X \mid Y)\}$ [By Double-E Rule] ...(2)

We replace Z by $E(X \mid Y)$ in (A) to obtain

$Var\{E(X \mid Y)\} = E\{E^2(X \mid Y)\} - [E \mid E(X \mid Y)]^2 = E\{E^2(X \mid Y)\} - E^2(X)$ [By Double-E Rule] ...(3)

Adding (2) and (3) : $E\{Var(X \mid Y)\} + Var\{E(X \mid Y)\} = E(X^2) - E^2(X) = Var(X)$.

7-31. *Formula for Conditional Covariance*

$Cov(X, Y) = E[Cov(X, Y) \mid T] + Cov[E(X \mid T), E(Y \mid T)]$ = mean (cond. cov) + cov (cond. means).

Proof. $\sigma_{XY} = Cov(X, Y) = E(XY) - E(X)\,E(Y)$ [Easy cutter] ...(1)

$= E[E(XY \mid T)\} - E\{E(X \mid T)\}\,E\{E(Y \mid T)\}$ [By Double-E Rule]

$= E[Cov[(X, Y) \mid T] + E(X \mid T)\,E(Y \mid T)] - E\{E(X \mid T)\}\,E\{E(Y \mid T)\}$. [by (1)]

$= E\{Cov[(X, Y \mid T)]\} + E\{E(X \mid T)\,.\,E(Y \mid T)\} - E[E(X \mid T)]\,[E\{E(Y \mid T)\}]$

$= E\{Cov[(X, Y) \mid T]\} + Cov[E(X \mid T), E(Y \mid T)]$, [By Easy Cutter]

Remarks. Setting $Y = X$, using $Cov(X, X) = Var(X)$, we recover the formula for the conditional variance §7-30.

7-32. *Correlation Ratio*

Definition. The correlation ratio of Y on X, written $\eta_{Y.X}$ is defined by

$$\eta_{Y.X} = \{Var(\mu_{Y \mid X})/\sigma_Y^2\}^{1/2},$$

Similarly, $\eta_{X.Y} = \{Var(\mu_{X \mid Y})/\sigma_X^2\}^{1/2}$ $[\mu_{Y \mid X} = E(Y \mid X)]$ ...(1)

Relation between correlation coefficient and correlation ratio :

$$|\rho_{XY}| \le \min\{\eta_{Y.X}, \eta_{X.Y}\} \qquad ...(2)$$

Proof. As $\rho_{XY} = \rho_{Y.X}$ it suffices to prove that $\rho^2_{X.Y} \le \mu^2_{Y.X}$.

Step 1. $Cov(X, \mu_{Y \mid X}) = Cov(X, Y)$ [§7-11(vii)]

Step 2. C-S Inequality : $[E(UV)]^2 \le E(U^2)\,.\,E(V^2)$. [§ 10-14] ...(3)

Take $U = X - u_X$, $V = \mu_{Y \mid X} - E(\mu_{Y \mid X}) = \mu_{Y \mid X} - \mu_Y$;

$\therefore \quad E(UV) = E\{X - \mu_X)\,.\,(\mu_{Y \mid X} - \mu_Y\} = Cov(X, \mu_{Y \mid X}) = Cov(X, Y) = \sigma_{XY}$.

$\therefore \quad \sigma^2_{XY} \le (\sigma^2_X)(\sigma^2_{\mu_{Y \mid X}})$

Step 3. Divide above by $\sigma^2_X\,.\,\sigma^2_Y$ to get

$$\rho^2_{XY} \le (\sigma^2_{\mu_{Y \mid X}})/\sigma^2_Y = \eta^2_{Y.X} \quad \Rightarrow \quad \rho_{XY} \le \eta_{Y \mid X}.$$

7-40. *Computations of Probabilities by Conditioning*

The Double E Rule $E(X) = E[E(X \mid Y)]$, in summation form is

$E(X) = \sum_y E(X \mid Y)\,P(Y = y)\,Y$: Discrete; $E(X) = \int_{-\infty}^{\infty} E(X \mid Y)\,f_Y(y)\,dy$ Y : Continuous ...(1)

Now, let A be an arbitrary event and define the indicator variate X by $X = 1$, if A occurs, and $X = 0$, if A does not occur. This provides, $E(X) = P(A)$ as well as $E(X \mid Y = y) = P(A \mid Y = y)$, for any variate Y. Hence (1) can be rewritten as

$$P(A) = \begin{cases} \Sigma_y P(A \mid Y = y) P(Y = y) & Y : \text{Discrete} \\ \int_{-\infty}^{\infty} P(A \mid Y = y) f_Y(y)\, dy & Y : \text{Continuous} \end{cases} \qquad \ldots(2)$$

This method is purely algebraic and does not require to consider geometric regions of the plane in order to obtain the limits of integration.

7-41. Double-E Rule, Cov $(X, \mu_{Y \mid x})$ and Tower Property

(i) $E(X) = E[E(X \mid Y)]$, (ii) $E[E(X \mid Y, Z) \mid Y] = E(X \mid Y)$.

Proof. Assume that X and Y are jointly continuous. Then using definition of $E(X)$,

$$E(X) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x \,.\, f(x, y)\, dy\, dx = \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} x \,.\, \frac{f(x, y)}{f_Y(y)}\, dx\right] f_Y(y)\, dy$$

$$= \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} x f_1(x \mid y)\, dx\right] f_Y(y)\, dy = \int_{-\infty}^{\infty} [E(X \mid y)]\, f_2(y)\, dy = E[E(X \mid Y)].$$

Aliter. Steps in the above proof can be reversed also :

$$E(X) = \int_{-\infty}^{\infty} x f_1(x)\, dx = \int_{-\infty}^{\infty} x \left[\int_{-\infty}^{\infty} f(x, y)\, dy\right] dx = \int_{-\infty}^{\infty} x \left\{\int_{-\infty}^{\infty} f(x \mid y)\, f_2(y)\, dy\right\} dx$$

$$= \int_{-\infty}^{\infty} \left\{\int_{-\infty}^{\infty} x f(x \mid y)\, dx\right\} f_2(y)\, dy = \int_{-\infty}^{\infty} \{E(X \mid Y)\}\, f_2(y)\, dy = \int_{-\infty}^{\infty} \phi(y) \,.\, f_2(y)\, dy$$

$$= E\{\phi(Y)\} = E\{E(X \mid Y)\}.$$

(ii) $\text{Cov}\,(X, \mu_{Y \mid X}) = E\{X \,.\, E(Y \mid X)\} - E(X) \,.\, E\{E(Y \mid X)\}$ [Def. of Cov]

$= E\{E(X \,.\, Y \mid X)\} - E(X) \,.\, E(Y)$ [Pull-through and Double-E Rules]

$= E(XY) - E(X)\, E(Y) = \text{Cov}\,(X, Y)$. [Double-E Rule]

(iii) Let X, Y, Z be continuous variates with joint p.d.f. $f(x, y, z)$. We shall not use subscripts for clarity, and arguments involved will be sufficient to specify the nature of density function. Now by definition.

$$E(X \mid Y, Z) = \int_{-\infty}^{\infty} x f(x \mid y, z)\, dx = \int_{-\infty}^{\infty} \frac{x f(x, y, z)}{f(y, z)}\, dx = g(Y, Z), \text{ say}$$

$$\therefore \quad E\{E(X \mid Y, Z) \mid Y\} = E[g(Y, Z) \mid Y]$$

$$= \int_{-\infty}^{\infty} f(z \mid y)\, g(y, z)\, dz = \int_{-\infty}^{\infty} \frac{f(y, z)}{f_Y(y)} \left[\int_{-\infty}^{\infty} \frac{x f(x, y, z)}{f(y, z)}\, dx\right] dz = \int_{-\infty}^{\infty} \frac{1}{f(y)} \left[\int_{-\infty}^{\infty} x f(x, y, z)\, dx\right] dz$$

$$= \int_{-\infty}^{\infty} \frac{x}{f(y)} \left[\int_{-\infty}^{\infty} f(x, y, z)\, dz\right] dx \qquad \text{(Assuming change of order in integration)}$$

$$= \int_{-\infty}^{\infty} \frac{x}{f(y)}\, f(x, y)\, dx = \int_{-\infty}^{\infty} x f(x \mid y)\, dx = E\,(X \mid Y).$$

7-50. Worked-out Problems

***Example* 1.** The joint p.m.f. of X and Y, $p(x, y) \equiv p_{xy}$ is $p_{11} = \frac{1}{3}$, $p_{12} = \frac{1}{9}$, $p_{13} = 0$, $p_{21} = \frac{1}{3}$, $p_{22} = 0$, $p_{23} \equiv \frac{1}{6}$, $p_{31} = \frac{1}{9}$, $p_{32} = \frac{1}{18}$, $p_{33} = \frac{1}{9}$.

Compute E$[X \mid Y = j]$, for $j = 1, 2, 3$.

$x \downarrow$ \ $y \rightarrow$	1	2	3	$f_1(x)$
1	1/9	1/9	0	2/9
2	1/3	0	1/6	1/2
3	1/9	1/18	1/9	5/18
$f_2(y)$	5/9	3/18	5/18	1

Solution. The adjoining table shows the given data and the marginal p.m.f.'s. Now

$$f(x \mid y) = \frac{f(x, y)}{f_2(y)}, \mathrm{E}(X \mid Y = y) = \frac{\Sigma\, xf(x, y)}{f_2(y)} \quad \ldots(1)$$

Now consider $y = 1, 2, 3$ and take note of columns 1, 2, 3 of the table :

$$\mathrm{E}(X \mid y = 1) = \tfrac{9}{5} \Sigma\, xf(x, 1) = \tfrac{9}{5}\left[1.\tfrac{1}{9} + 2.\tfrac{1}{3} + 3.\tfrac{1}{9}\right] = 2$$

$$\mathrm{E}(X \mid y = 2) = \tfrac{18}{3} \Sigma\, xf(x, 2) = 6\left[1.\tfrac{1}{9} + 2.0 + 3.\tfrac{1}{18}\right] = \tfrac{5}{3}$$

$$\mathrm{E}(X \mid y = 3) = \tfrac{18}{5} \Sigma\, xf(x, 3) = \tfrac{18}{5}\left[1.0 + 2.\tfrac{1}{6} + 3.\tfrac{1}{9}\right] = \tfrac{12}{5}.$$

***Example* 2.** The life times of inverters are identically distributed with c.d.f. and p.d.f. as F and f respectively. In terms of F, f and θ, find expected value of the life time of a θ-hour-old inverter.

Solution. Let X be the life-time of the θ-hour-old inverter. We need E$(X \mid X > \theta)$. Write $A = \{X > \theta\}$. Now

$$F_{X \mid A}(t) = P\{X \le t \mid A\}, \quad f_{X \mid A}(t) = F'_{X \mid A}(t)$$

So
$$F_{X \mid A}(t) = \frac{P(X \le t, A)}{P(A)} = \frac{P(X \le t, X > \theta)}{P(X > \theta)} = \frac{P\{\theta < X \le t\}}{P\{X > \theta\}} = \frac{F(t) - F(\theta)}{1 - F(\theta)}.$$

This evaluation holds if $t > \theta$ and equals zero if $t \le \theta$.

Differentiation yields,

$$f_{X \mid A}(t) = \frac{f(t)}{[1 - F(\theta)]} \quad \text{if} \quad t > \theta$$
$$= 0, \quad \text{if } t \le \theta.$$

$$\therefore \quad \mathrm{E}(X \mid X > \theta) = \int_0^{\infty} t \,.\, f_{X \mid A}(t)\, dt = \frac{1}{1 - F(\theta)} \cdot \int_{\theta}^{\infty} t\, f(t)\, dt.$$

Comment. If inverter, installed at $t = 0$, begins to operate, gets dead at $t = \theta$, Then E $(X \mid X < \theta)$, expected life-time of the dead-inverter, is

$$\mathrm{E}(X \mid X < \theta) = \frac{1}{F(\theta)} \int_0^{\theta} tf(t)\, dt.$$

***Example* 3.** Independent trials, each of which with success probability p, are performed until there are k consecutive successes. Find the mean number of necessary trials.

Solution. Write N_k for the number of necessary trials to obtain k consecutive successes. Let $M_K = \mathsf{E}(N_k)$. We find recursive relation between M_k and M_{k-1} by using Double-E Rule. Thus

$$M_k = \mathsf{E}(N_k) = \mathsf{E}\{\mathsf{E}(N_k \mid N_{k-1})\} \quad \text{...(1)}$$

Now, $$\mathsf{E}(N_k \mid N_{k-1}) = \{N_{k-1} + 1\} + q\,\mathsf{E}\{N_k\}. \quad \text{...(2)}$$

Relation (2) follows, because, if it takes N_{k-1} trials to produce $(k-1)$ consecutive successes, then either the next trial is a success, whence we get k consecutive success or the next trial is a failure (probability $q = 1 - p$) and we must start a fresh for objective N_k. Now taking expectation of both sides of (2) yields.

$$M_k = \mathsf{E}\{N_{k-1} + 1 + q\mathsf{E}(N_k)\} \equiv M_{k-1} + 1 + q \,.\, M_k.$$

$$\Rightarrow \qquad M_k = \left(\frac{1}{p}\right) + \left(\frac{1}{p}\right) M_{k-1} \quad \text{...(3)}$$

Since N_1 = the time of first success is gem(p), with $\mathsf{E}(N_1) = 1/p$, *i.e.* $M_1 = 1/p$, the recursion (3) gives

$$M_2 = \left(\frac{1}{p}\right) + \left(\frac{1}{p^2}\right), \qquad M_3 = \left(\frac{1}{p}\right) + \left(\frac{1}{p}\right)^2 + \left(\frac{1}{p}\right)^3$$

and hence in general, (or by induction process)

$$M_k = \left(\frac{1}{p}\right) + \left(\frac{1}{p}\right)^2 + \left(\frac{1}{p}\right)^3 + \ldots + \left(\frac{1}{p}\right)^k = \frac{(1-p^k)}{(1-p)\,p^k}.$$

***Example* 4.** Variates X and Y take the values 1, 2, 3 along with the probabilities shown. Find p, $\mathsf{E}(X + Y)$, Var (X), Cov (X, Y). Determine whether X and Y are independent. Given $Y = 2$, what is the conditional probability distribution of X ? Find also Var $(Y \mid X = 1)$.

$x\downarrow$ \ $y\rightarrow$	1	2	3
1	p	p	p
2	p	$2p$	p
3	p	p	p

Solution. Since $\sum f(x, y) = 1$ (Normality), we must have $10p = 1$, so $p = 0.1$.

Now row totals and column totals represent the absolute (marginal) probabilities $P(X = a)$ and $P(Y = b)$, etc. Further X and Y are symmetric distributed. Now

$\mathsf{E}(X) = 1(3p) + 2(4p) + 3(3p) = 20p = 2$, $\mathsf{E}(Y) = \mathsf{E}(X)$, so $\mathsf{E}(X + Y) = \mathsf{E}(X) + \mathsf{E}(Y) = 2\mathsf{E}(X) = 4$.

$\mathsf{E}(X)^2 = 1^2 \,.\, 3p + 2^2 \,.\, 4p + 3^2 \,.\, 3p = 3p + 16p + 27p = 46p = \mathsf{E}(Y^2)$.

$$\text{Var }(X) = \mathsf{E}(X^2) - \mathsf{E}^2(X) = 46p - 4 = 4.6 - 4 = 0.6 = \text{Var }(Y).$$

$$\mathsf{E}(XY) = (p + 2p + 3p) + (2p + 8p + 6p) + (3p + 6p + 9p) = 40p = 4.$$

$$\text{Cov }(X, Y) = \mathsf{E}(XY) - \mathsf{E}(X)\,\mathsf{E}(Y) = 4 - 2 \times 2 = 0.$$

It follows that X and Y are uncorrelated. However, X and Y are not independent for,

$$P(X = 3, Y = 3) = p \neq P(X = 3)\,P(Y = 3) = 9p^2.$$

Conditional distributions. Obviously, $P(Y = 2) = 4p$.

$P(X = x \mid Y = 2) = P(x, 2)/P(Y = 2) = P(x, 2)/4p$; $P(X = 1 \mid Y = 2) = P(1, 2)/4p = p/4p = \frac{1}{4}$.

$P(X = 2 \mid Y = 2) = P(2, 2)/4p = 2p/4p = \frac{1}{2}$, $P(X = 3 \mid Y = 2) = P(3, 2)/4p = p/4p = \frac{1}{4}$.

Further, $$E(Y \mid X = 1) = \sum_y y f(y \mid x = 1) = \sum_y \frac{y f(1, y)}{P(X = 1)} = \frac{\Sigma y f(1, y)}{3p}$$

$$E(Y \mid X = 1) = \frac{(p + 2p + 3p)}{3p} = 2. \quad E(Y^2 \mid X = 1) = \frac{\Sigma y^2 f(1, y)}{3p} = \frac{(p + 4p + 9p)}{3p} = \frac{14}{3}.$$

$$\text{Var}(Y \mid X = 1) = E(Y^2 \mid X = 1) - E^2(Y \mid X = 1) = (14/3) - 4 = \frac{2}{3}.$$

Example 5. Let $f(x, y) = 24xy$, $x > 0$, $y > 0$, $x + y \le 1$, $f(x, y) = 0$, elsewhere. Show that $\text{Var}(Y \mid x) = (x - 1)^2/18$.

Solution. Here : $$f_X(x) = \int_0^{1-x} 24xy \, dy = 12x(1 - x)^2, \quad 0 < x < 1$$

$$\therefore \quad f(y \mid x) = \frac{f(x, y)}{f_X(x)} = \frac{2y}{(1 - x)^2}, 0 < y < 1 - x$$

Now $$E(Y^k \mid x) = \int_0^{1-x} y^k f(y \mid x) \, dy = \frac{2}{(1 - x)^2} \int_0^{1-x} y^{(k+1)} \, dy = \frac{2(1 - x)^k}{(k + 2)}.$$

Thus, $$E(Y \mid x) = \frac{2(1 - x)}{3}; \quad E(Y^2 \mid x) = \frac{(1 - x)^2}{2}.$$

$$\therefore \quad \text{Var}(Y \mid x) = E(Y^2 \mid x) - [E(Y \mid x)]^2 = \tfrac{1}{2}(1 - x)^2 - \tfrac{4}{9}(1 - x)^2 = \tfrac{1}{18}(x - 1)^2.$$

Example 6. The joint p.d.f. of X, Y is $f(x, y) = Cxy(2 - x - y)$, $0 < x, y < 1$. Evaluate $E(X \mid Y = y)$.

Solution. We determine C by normality $\Sigma\Sigma f(x, y) = 1$. Now

$$\sum_x f(x, y) = Cy \int_0^1 (2x - x^2 - xy) \, \partial x = Cy \frac{(4 - 3y)}{6} \quad \text{...(i)}$$

$$\therefore \quad \sum_y \sum_x f(x, y) = \frac{C}{6} \int_0^1 (4y - 3y^2) \, dy = \frac{C}{6}.$$

Thus, by normality, $C/6 = 1 \Rightarrow C = 6$ and (i) thus yields

$$f_2(y) = \sum_x f(x, y) = y(4 - 3y), \quad 0 < y < 1. \quad \text{...(ii)}$$

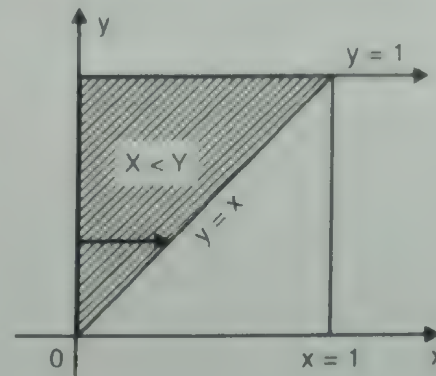
Consequently, $$f(x \mid y) = \frac{f(x, y)}{f_2(y)} = \frac{6x(2 - x - y)}{(4 - 3y)}$$

So $$E(X \mid y) = \frac{6}{4 - 3y} \int_0^1 x^2 (2 - x - y) \, \partial x$$

$$= \frac{6}{4 - 3y} \left(\frac{2}{3} - \frac{1}{4} - \frac{1}{3} y \right) = \frac{5 - 4y}{2(4 - 3y)}, \quad 0 < y < 1.$$

Example 7. Let $f(x, y) = 21\,x^2\,y^3$, $0 < x < y < 1$; $f(x, y) = 0$, elsewhere ; be the joint p.d.f. of X and Y. Find the conditional mean and variance of X, given $Y = y$, $0 < y < 1$.

Solution. We determine the marginal density $f_Y(y)$



$$f_Y f(y) = \int_0^y f(x, y)\,dx = 21y^3 \int_0^y x^2\,dx = 7y^6,\quad 0 < y < 1.$$

$$\therefore \quad f(x \mid y) = \frac{f(x, y)}{f_Y(y)} = \frac{3x^2}{y^3}.$$

$$\mathrm{E}(X^k \mid y) = \int_0^y x^k \cdot f(x \mid y)\,dx = \frac{3}{y^3}\int_0^y x^{k+2}\,dx = \frac{3y^k}{k+3}.$$

$$\therefore \quad \mathrm{E}(X \mid y) = \frac{3y}{4},\ \mathrm{E}(X^2 \mid 4) = \frac{3y^2}{5}.$$

$$\therefore \quad \mathrm{Var}\,(X \mid y) = \mathrm{E}(X^2 \mid y) - \mathrm{E}^2(X \mid y) = \frac{3}{5}y^2 - \frac{9}{16}y^2 = \frac{3}{80}y^2.$$

Example 8. Compute $\mathrm{E}(Y \mid X = x)$ and $\mathrm{E}(Y)$ for the joint density

$f(x, y) = n(n-1)\,(y-x)^{n-2}$, $0 \le x \le y \le 1$; $f(x, y) = 0$, elsewhere.

Solution. Marginal density $f_1(x) = \int_{-\infty}^{\infty} f(x, y)\,dy = n(n-1)\int_x^1 (y-x)^{n-2}\,dy = n(1-x)^{n-1}, 0 \le x \le 1.$

$$\therefore\ f(y \mid x) = \frac{f(x, y)}{f_1(x)} = \frac{(n-1)\,(y-x)^{n-2}}{(1-x)^{n-1}},\quad x \le y \le 1.$$

$$\mathrm{E}(Y \mid x) = \frac{(n-1)}{(1-x)^{n-1}}\int_x^1 y(y-x)^{n-2}\,dy = (n-1)\int_0^1 t^{n-2}[x + (1-x)t]\,dt. \qquad [y - x = (1-x)t]$$

$$= (n-1)\left[\frac{x}{n-1} + \frac{1-x}{n}\right] = \frac{n-1+x}{n}. \qquad \text{...(1)}$$

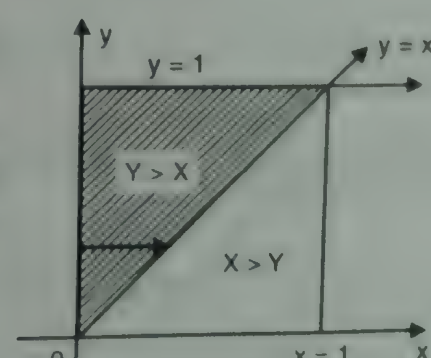
$$\mathrm{E}(X) = n\int_0^1 x(1-x)^{n-1}\,dx = n\,.\,B(2, n) = \frac{1}{n+1}.$$

$$\therefore \quad \mathrm{E}(Y) = \mathrm{EE}(Y \mid X) + \frac{n-1+\mathrm{E}(X)}{n} = \frac{n}{n+1}.$$

Example 9. Let $f(x, y) = 8xy$, $0 < x < y < 1$; $f(x, y) = 0$, elsewhere.

Find (a) $\mathrm{E}(Y \mid X = x)$, (b) $\mathrm{E}(XY \mid X = x)$, (c) $\mathrm{Var}\,(Y \mid X = x)$.

Solution. Firstly, we determine marginal and conditional distributions : [refer to Figure above in Example 7]



$$f_X(x) = \int_{-\infty}^{\infty} f(x, y)\,dy = 8x\int_x^1 y\,dy = 4x(1-x^2),\quad 0 \le x \le 1.$$

$$f_Y(y) = \int_{-\infty}^{\infty} f(x, y)\,dx = 8y\int_0^y x\,dx = 4y^3,\quad 0 \le y \le 1.$$

$$f_1(x \mid y) = \frac{f(x, y)}{f_Y(y)} = \frac{2x}{y^2},\ 0 \le x \le y \le 1;\ f_2(y \mid x) = \frac{f(x, y)}{f_X(x)} = \frac{2y}{1 - x^2},\ 0 \le x \le y \le 1$$

Since $f(x, y) \ne f_X(x)\, f_Y(y)$, X and Y are not independent. Further

$$\mathrm{E}(Y^k \mid x) = \int_x^1 y^k \cdot \left(\frac{2y}{1 - x^2}\right) dy = \frac{k}{k+2}\left(\frac{1 - x^{k+2}}{1 - x^2}\right)$$

(a) $\therefore$ $$\mathrm{E}(Y \mid x) = \frac{2}{3}\left(\frac{1 - x^3}{1 - x^2}\right) = \frac{2}{3}\left(\frac{1 + x + x^2}{1 + x}\right)\mathrm{E}(Y^2 \mid X = x) = \frac{1 + x^2}{2}.$$

(b) $$\mathrm{E}(X_{Y \mid X} = x) = x\mathrm{E}(Y \mid x) = \frac{2}{3}\,\frac{x(1 + x + x^2)}{(1 + x)}, \qquad \text{[by (a)]}$$

$$\mathrm{Var}\,(Y \mid x) = \mathrm{E}(Y^2 \mid x) - \mathrm{E}^2(Y \mid x) = \frac{1 + x^2}{2} - \frac{4}{9}\,\frac{(1 + x + x^2)^2}{(1 + x)^2} = \frac{(x - 1)^2\,(x^2 + 4x + 1)}{18(1 + x)^2}.$$

***Example* 10.** A variate X has a p.d.f. $f(x) = 2x/a^2$, $0 < x < a$. Variate Y has p.m.f.

$$P(y) = P\{(ay/10) < X < a(y + 1)/10\}; \qquad (y = 0, 1, 2, \ldots, 9).$$

Find $\mathrm{E}(Y)$ and $P(Y = 8 \mid Y > 7\}$.

Calculate $\mathrm{E}(e^{tX} \mid X < a/2)$ and hence find $\mathrm{E}(X \mid X < a/2)$.

Solution. $$P_Y(y) = \int_{ay/10}^{a(y+1)/10} \frac{2x}{a^2}\,dx = \frac{1}{a^2}\left[\frac{a^2(1 + y)^2}{100} - \frac{a^2 y^2}{100}\right] = \frac{2y + 1}{100}.$$

$$\mathrm{E}(Y) = \sum_{y=0}^{9}\left(\frac{2y + 1}{100}\right)y = \frac{1}{100}\sum_{y=0}^{9}(2y^2 + y) = \frac{1}{100}\left\{\frac{9.10.19}{3} + \frac{9.10}{2}\right\} = \frac{123}{20} = 6.15.$$

$$P(Y = 8 \mid Y > 7) = \frac{P(Y = 8)}{P(Y > 7)} = \frac{17/100}{(17/100) + (19/100)} = \frac{17}{36}$$

$$\text{Now } p = P\left(X < \frac{a}{2}\right)\int_0^{a/2} \frac{2x}{a^2}\,dx = \frac{1}{4} \qquad \ldots(1)$$

$$\mathrm{E}\left(e^{tX} \mid X < \frac{a}{2}\right) = \int_0^{a/2} e^{tX} \cdot \frac{f(x)}{P(X < a/2)}\,dx = 4\int_0^{a/2} e^{tX}\,\frac{2x}{a^2}\,dx, \qquad \text{[by (1)]}$$

$$= 8\left[1 + \left(\tfrac{1}{2}at - 1\right)e^{at/2}\right]/a^2 t^2 \qquad \ldots(2)$$

$$\mathrm{E}\left[X \mid X < \left(\frac{a}{2}\right)\right] = \text{Coefficient of } t \text{ in } \mathrm{E}\left(e^{tX} \mid x < \left(\frac{a}{2}\right)\right)$$

$$= \text{Coefficient of } t \text{ in } \frac{8}{a^2 t^2}\left[1 + \left(\frac{1}{2}at - 1\right)\left(1 + \frac{at}{2} + \frac{a^2 t^2}{8} + \frac{a^3 t^3}{48} + \ldots\right)\right] = \frac{a}{3}.$$

Note. Expression (2) is really m.g.f. of X, conditioned on the region $X < 1/2a$. Concept of m.f.g. is fully exploited in Chapter 8.

***Example* 11.** Three coins are tossed. Let X denote the number of heads on the first two coins, Y the number of tails on the last two and z denote the number of heads on the last two.

(a) Find the joint distribution of (i) X and Y, (ii) X and Z

(b) Find the Cond. distribution of Y given $X = 1$.

(c) Find $E(Z \mid X = 1)$,   (d) $\rho_{X, Y}$ and $\rho_{X, Z}$.

(e) Give a joint distribution that is not the joint distribution of X and Z in (a) yet has the same marginals as $f(x, z)$ has in part (a).

Solution. We enumerate the values of variates and note that probability $p = 1/8$ for each outcome ω, the sample point.

ω	x	y	z	p
hhh	2	0	2	p
hht	2	1	1	p
hth	1	1	1	p
thh	1	0	2	p

ω	x	y	z	p
htt	1	2	0	p
tht	1	1	1	p
tth	0	1	1	p
ttt	0	2	0	p

Joint Distribution

$x\downarrow$ \ $y\rightarrow$	0	1	2	$P(x)$
0	0	p	p	$2p$
1	p	$2p$	p	$4p$
2	p	p	0	$2p$
$P(y)$	$2p$	$4p$	$2p$	$8p = 1$

$x\downarrow$ \ $z\rightarrow$	0	1	2	$P(x)$
0	p	p	0	$2p$
1	p	$2p$	p	$4p$
2	0	p	p	$2p$
$P(z)$	$2p$	$4p$	$2p$	$8p = 1$

(a) The joint distributions of (X, Y) and (X, Z) are indicated in above tables.

(b)
$$p_1 = P(X=1) = 4p = \frac{1}{2}.$$

$$P(Y=2 \mid X=1) = \frac{P(1,2)}{p_1} = \frac{p}{4p} = \frac{1}{4}.$$

$$P(Y=1 \mid X=1) = \frac{P(1,1)}{p_1} = \frac{2p}{4p} = \frac{1}{2};\ P(Y=0 \mid X=1) = \frac{P(1,0)}{p_1} = \frac{p}{4p} = \frac{1}{4}.$$

(c)
$$P(Z=2 \mid X=1) = \frac{P(1,2)}{p_1} = \frac{p}{4p} = \frac{1}{4}.$$

$$P(Z=1 \mid X=1) = \frac{P(1,1)}{p_1} = \frac{2p}{4p} = \frac{1}{2}.\ P(Z=0 \mid X=1) = \frac{P(1,0)}{p_1} = \frac{p}{4p} = \frac{1}{4}.$$

$$\therefore \quad E(Z \mid X=1) = \Sigma z\,.\,P(z \mid X=1) = 2\left(\frac{1}{4}\right)+1\left(\frac{1}{2}\right)+0\left(\frac{1}{4}\right)=1.$$

(d) $E(X) = 4p+4p+0=1=E(Y);\ E(X^2)=8p+4p=\frac{3}{2}=E(Y^2).\,\sigma_X^2=\sigma_Y^2=\sigma_Z^2=\frac{1}{2}$

$$E(XY) = 2p+4p=\frac{3}{4}, \text{Cov}(X,Y)=\left(\frac{3}{4}\right)-1=-\frac{1}{4}.\ \text{Corr}(X,Y)=-\frac{1}{2};$$

$$E(XZ) = 10p=\frac{5}{4}, \ \text{Cov}(X,Z)=\left(\frac{5}{4}\right)-1=\frac{1}{4}.\ \text{Corr}(X,Z)=\frac{1}{2}.$$

(e) Let $0 \le \varepsilon \le p$. The joint distribution of (X, Z) shown in the following table has the same marginals as in part (a). Another possible solution is

$z \rightarrow$ / $x \downarrow$	0	1	2	$P(x)$
0	$p+\varepsilon$	$p-\varepsilon$	0	$2p$
1	$p-\varepsilon$	$2p+\varepsilon$	p	$4p$
2	0	p	p	$2p$
$P(z)$	$2p$	$4p$	$2p$	1

$$f_1(x,z) = f_1(x), f_3(z)$$

$$f_1(x) : (2, 2p), (1, 4p), (0, 2p)$$

$$f_3(z) : (2, 2p), (1, 4p), (0, 2p).$$

***Example* 12.** A fair die is cast 3 times. The r.v. X is the number of occurrence of a six and the r.v. Y is the number of occurrence of an odd number. Find the joint distribution of (X, Y), $P(X > Y)$ and $P(X + Y \ge 2)$. Prove that $E(X \mid Y = 2) = 1/3$ and Corr $(X, Y) = -1/\sqrt{5}$.

***Solution. Step* 1.** The simplest way to find the joint distribution p_{xy}, in this case, is to first find the individual densities of X and Y, which we proceed to obtain.

The event $\{X = i\}$, $0 \le i \le 3$, is a set of outcomes from Ω such that there occur exactly i sixes in them. Here, $P(X = i)$ means, prob. of getting exactly i sixes in 3 Bernouli trials with success-prob. in one trial being $p = 1/6$. Thus

$$P(X=i) = \binom{3}{i} q^{3-i} p^i = \binom{3}{i}\left(\tfrac{5}{6}\right)^{3-i}\left(\tfrac{1}{6}\right)^i,\ i=0,1,2,3. \qquad ...(1)$$

Since P {odd eye on die} = 1/2 = p (say), the distribution of Y is

$$P\{Y=j\} = \binom{3}{j} q^{3-j} p^j = \binom{3}{j}\left(\tfrac{1}{2}\right)^3 = \left(\tfrac{1}{8}\right)\binom{3}{j},\ j=0,1,2,3. \qquad ...(2)$$

We now fill the row totals $[P(x_i)]$ and the column totals $[P(y_j)]$ of the bivariate (matrix) table.

***Step* 2.** Some entries in the prob. table correspond to *impossible events*, which yield

$$p_{31} = p_{32} = p_{33} = p_{13} = p_{23} = p_{22} = 0.$$

Note that, e.g. $p_{23} = P\{2$ sixes and 3 odd numbers$\} = P(\varnothing) = 0$. Others follow similarly. Some entries are trivially obtained, e.g.

$$p_{30} = P(X=3, Y=0) = \left(\tfrac{1}{6}\right)^3,\ p_{03} = P(X=0, Y=3) = \left(\tfrac{1}{2}\right)^3 = 8.$$

***Step* 3.** Some entries need *auxiliary* evaluations. Thus

$$p_{20} = P(X = 2, Y = 0) = P(Y = 0)\ P(X = 2 \mid Y = 0) = (1/8)\ P(X = 2 \mid Y = 0).$$

Now $P(X = 2 \mid Y = 0) = P$ (draw 2 sixes from three items $\{2, 4, 6\}$, with replacement and ordering) $= \binom{3}{2}\left(\frac{2}{3}\right)^{3-2}\left(\frac{2}{3}\right)^{2} = \frac{2}{9}.$

$x_i \backslash y_i$	0	1	2	3	$P(x_i)$
0	1/27	1/6	1/4	1/8	125/216
1	1/18	1/6	1/8	0	25/72
2	1/36	1/24	0	0	5/72
3	1/216	0	0	0	1/216

Thus, $\quad p_{20} = \frac{2}{72} = \frac{1}{36}.$

Likewise, $p_{10} = P(Y = 0)\ P(X = 1 \mid Y = 0)$

$$= \frac{1}{8}\cdot\binom{3}{1}\left(\frac{2}{3}\right)^{2}\left(\frac{1}{3}\right) = \frac{1}{18}.$$

Now $\quad p_{01} = P(X = 0)\ P(Y = 1 \mid X = 0) = \frac{125}{216}\ P(Y = 1 \mid X = 0).$

$$P(Y = 1 \mid X = 0) = P \text{ (one odd number from } \{1, 2, 3, 4, 5\})$$

$$= \binom{3}{1} q^2 p = \binom{3}{1}\left(\tfrac{2}{5}\right)^2\left(\tfrac{3}{5}\right) = \frac{36}{125}$$

$$\therefore \qquad p_{01} = \frac{125}{216}\cdot\frac{36}{125} = \frac{1}{6}.$$

All the remaining probabilities (entries) can be had from simple arrangements of marginal totals. Simple subtractions in the order : $p_{21}, p_{11}, p_{12}, p_{02}, p_{00}$ yield all. Thus

$$p_{21} = \frac{5}{72} - \frac{1}{36} = \frac{1}{24},\ p_{11} = \frac{3}{8} - \frac{1}{6} - \frac{1}{24} = \frac{1}{6},\ p_{12} = \frac{25}{72} - \frac{1}{18} - \frac{1}{6} = \frac{1}{8},\ p_{02} = \frac{3}{8} - \frac{1}{8} = \frac{1}{4},\ p_{00} = \frac{1}{27}.$$

We note that $P(X = 1, Y = 1) \neq P(X = 1)\ P(Y = 1)$, whence X and Y are not independent.

$P(X \geq Y) =$ Sum of prob. on and below main diagonal $= \frac{1}{8} + \frac{5}{24} = \frac{1}{3}.$

$$P(X + Y \geq 2) = 1 - P(X + Y \leq 1) = 1 - [p_{00} + p_{01} + p_{10}] = 20/27.$$

Second Part. $P(X = i \mid Y = 2) = P(i, 2)/P(Y = 2) = \frac{8}{3}\ p_{i2},\ i = 0, 1, 2, 3.$

$$\therefore\ E(X \mid Y = 2) = 0.\frac{2}{3} + 1.\frac{1}{3} + 2.0 + 3.0 = \frac{1}{3}.\ E(X) = 0 + \frac{25}{72} + \frac{10}{72} + \frac{1}{72} = \frac{1}{2},\ E(X^2) = 0 + \frac{25}{72} + \frac{20}{72} + \frac{3}{72} = \frac{2}{3}$$

$$E(Y) = 0 + \tfrac{3}{8} + \tfrac{6}{8} + \tfrac{3}{8} = \tfrac{3}{2},\ E(Y^2) = 0 + \tfrac{3}{8} + \tfrac{12}{8} + \tfrac{1}{8} = 3,\ E(XY) = \tfrac{1}{6} + \tfrac{1}{4} + \tfrac{1}{12} = \tfrac{1}{2}.$$

$\sigma_{XY} = E(XY) - E(X)\ E(Y) = -1/4.\ \sigma_X^2 = E(X^2) - E^2(X) = 5/12,\ \sigma_X^2 = E(Y^2) - E^2(Y) = \frac{3}{4}.$

$$\therefore \qquad \rho_{XY} = \frac{\sigma_{XY}}{\sigma_X\,\sigma_Y} = \frac{-1/4}{\sqrt{5/12}\,\sqrt{3/4}} = -\frac{1}{\sqrt{5}} = -\frac{\sqrt{5}}{5} = -0.45.$$

***Example* 13.** Show that if Z and X are dependent variates, Z can be independent of X given Y.

Solution. Consider $f(x, y, z) \equiv P(X = x, Y = y, Z = z) = p^3 q^{z-3}\ ;\ 0 < p\ ,\ q < 1,\ p + q = 1.$

$$x = 1, 2, ..., y-1, \;\; y = 2, 3, ..., z-1, \; z = 3, 4, 5, ...$$

$$\therefore \quad f(x, y) = p^3 q^{-3} \sum_{z=y+1}^{\infty} q^z = p^3 q^{-3} q^{y+1}(1+q+q^2+...) = p^2 q^{y-2}, \; 1 \le x \le y-1, \; y = 2, 3, ...$$

$$f(y, z) = p^3 q^{z-3} \sum_{x=1}^{y-1} (1) = (y-1)\, p^3 q^{z-3}, \qquad \text{with above support.}$$

$$f(z, x) = p^3 q^{z-3} \sum_{y=x+1}^{z-1} (1) = (z-x+1)\, p^3 q^{z-3}, \; 1 \le x \le z-2; \; z = 3, 4, ...$$

$$f_1(x) = \sum_{y=x+1}^{\infty} f(x, y) = p^2 q^{-2}(q^{x+1} + q^{x+2} + ...) = pq^{x-1}, \; x = 1, 2, 3, ...$$

$$f_2(y) = \sum_{x=1}^{y-1} f(x, y) = (y-1)\, p^2 q^{y-2}, \; y = 2, 3, ..., \infty$$

$$f_3(z) = \sum_{y=2}^{z-1} f(y, z) = p^3 q^{z-3} \sum_{y=2}^{z-1} (y-1) = p^3 q^{z-3}[1+2+...+(z-2)] = \tfrac{1}{2}(z-1)(z-2)\, p^3 q^{z-3}.$$

Observe that $f(x, z) \ne f_1(x) f_3(z)$; $f(x, y) \ne f_1(x) f_2(y)$, $f(y, z) \ne f_2(y) f_3(z)$.

We conclude : X and Y dependent, Y and Z dependent, Z and X dependent

$$f(z \mid x, y) = f(x, y, z)/f(x, y) = pq^{z-y-1}, \; x = 1, 2, ..., y-1 \; ; \; y = 2, 3, ..., z-1.$$

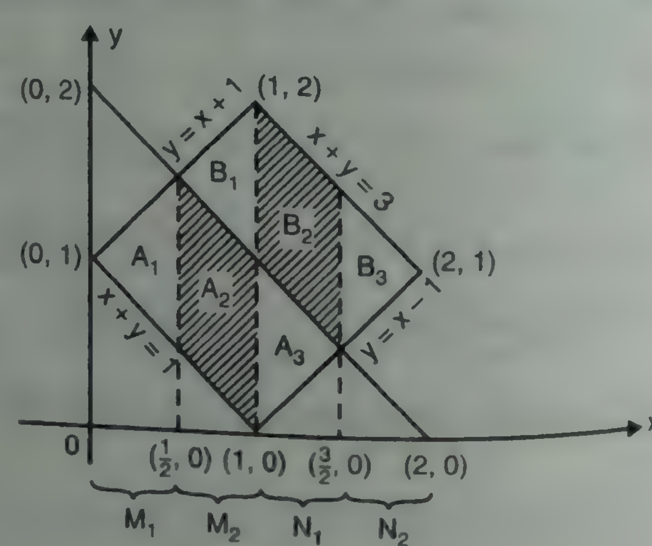
$$\mathrm{E}(Z \mid x, y) = \sum_{z=y+1}^{\infty} zpq^{z-y-1} = p \sum_{t=1}^{\infty} (y+t)\, q^{t-1} \qquad [t = z-y]$$

$$= py\Sigma q^{t-1} + p\,\Sigma t\, q^{t-1} = py(1-q)^{-1} + p(1-q)^{-2} = y + p^{-1}$$

Thus, $\quad \mathrm{E}(Z \mid X, Y) = Y + p^{-1}$ a.s. (independent of X).

Conclusion. Z is conditionally indep. of X given Y despite the fact that Z and X are dep.

***Example* 14.** The pair (X, Y) has joint p.d.f. $f(x, y) = \frac{1}{2}\, xy$ on the square S with vertices $(1, 0)$, $(2, 1)$, $(1, 2)$, $(0, 1)$. Determine $\mathrm{E}(Z^k \mid X = x)$ where $Z = X$, when $X + Y = 2$, $Z = X^2 Y$, when $X + Y > 2$.



Solution. We designate the regions and the intervals indicated in the figure by

$I_U = \{(x, y) : (x + y > 2) \cap S\} = B_1 \uplus B_2 \uplus B_3.$

$I_L = \{(x, y) : (x + y \le 2) \cap S\} = A_1 \uplus A_2 \uplus A_3$

$M = [0, 1] = [0, \tfrac{1}{2}] \cup [\tfrac{1}{2}, 1] = M_1 \cup M_2$;

$N = N_1 \cup N_2 = [1 \le x \le \tfrac{3}{2}] \cup [\tfrac{3}{2} \le x \le 2] = [1, 2]$

$f_1(x) = \frac{1}{2}x\int_{1-x}^{1+x} y\,dy = x^2,\ 0 \le x \le 1;\ f_1(x) = \frac{1}{2}x\int_{x-1}^{3-x} y\,dy\ x(2-x),\ \ 1 \le x \le 2.$

$\therefore\ f_1(x) = x^2 I_M + x(2-x)\,I_N.\ f(y\,|\,x) = \dfrac{f(x,y)}{f_1(x)} = \dfrac{1}{2}\dfrac{y}{x}I_M + \dfrac{y}{2(2-x)}I_N.$

Now $Z = XI_L + X^2 Y I_U,\ \ Z^k = X^k I_L + X^{2k} Y^k I_U.$

$\mathsf{E}(Z^k\,|\,X=x) = \mathsf{E}\{(X^k I_L + X^{2k}Y^k I_U)\,|\,X=x\} = x^k\,\mathsf{E}(I_L\,|\,X=x) + x^{2k}\,\mathsf{E}(Y^k I_U\,|\,X=x)$, [by Lin E]

$$= x^k\int I_L f(y\,|\,x)\,dy + x^{2k}\int I_U y^k f(y\,|\,x)\,dy = x^k J_1 + x^{2k} J_1 \quad \text{(say)} \qquad \ldots(1)$$

$$J_1 = \int I_L\left[\frac{y}{2x}I_M + \frac{y}{2(2-x)}I_N\right]dy \qquad [I_L\cdot I_M = A_1 \uplus A_2,\ I_L\cdot I_N = A_3]$$

$$= \frac{1}{2x}\left\{\int_{A_1} y\,dy + \int_{A_2} y\,dy\right\} + \frac{1}{2(2-x)}\int_{A_3} y\,dy$$

$$= \frac{1}{4x}[(1+x)^2-(1-x^2)]_{M_1} + \frac{1}{4x}[(2-x)^2-(1-x)^2]_{M_2} + \frac{1}{4(2-x)}[(2-x)^2-(x-1)^2]_{N_1}$$

$$= [1]_{M_1} + \left[\frac{3-2x}{4x}\right]_{M_2} + \left[\frac{3-2x}{4(2-x)}\right]_{N_1} \qquad \ldots(2)$$

$$J_2 = \int I_U\left[\frac{y}{2x}I_M + \frac{y}{2(2-x)}I_N\right]y^k\,dy \qquad [I_U\cdot I_M = B_1,\quad I_U\cdot I_N = B_2 \uplus B_3]$$

$$= \frac{1}{2x}\int_{B_1} y^{k+1}\,dy + \frac{1}{2(2-x)}\left\{\int_{B_2} y^{k+1}\,dy + \int_{B_3} y^{k+1}\,dy\right\} \qquad [\text{Put } k+2 = \lambda]$$

$$= \frac{1}{2\lambda x}\left[(1+x)^\lambda-(2-x)^\lambda\right]_{M_2} + \frac{1}{2\lambda(2-x)}\left\{[(3-x)^\lambda-(2-x)^\lambda]_{N_1} + [(3-x)^\lambda-(x-1)^\lambda]_{N_2}\right\} \ldots(3)$$

Substituting from (2) and (3) into (1) yields

$$\mathsf{E}(Z^k\,|\,X=x) = x^k\left\{(1)_{M_1} + \left(\frac{3-2x}{4x}\right)_{M_2} + \left(\frac{3-2x}{4(2-x)}\right)_{N_1}\right\}$$

$$+\frac{x^{2k}}{2\lambda}\left\{\left(\frac{(1+x)^\lambda-(2-x)^\lambda}{x}\right)_{M_2} + \left(\frac{(3-x)^\lambda-(2-x)^\lambda}{(2-x)}\right)_{N_1} + \left(\frac{(3-x)^\lambda-(x-1)^\lambda}{(2-x)}\right)_{N_2}\right\} \ldots(4)$$

The indicator functions pick out the correct expressions for each value of x in any interval. Of course, $\mathsf{E}(Z^k\,|\,X=x) = 0$, for $x > 2$. The conditional mean $\mathsf{E}(Z\,|\,X=x)$ is had when $k = 1$. Thus

$$\mathsf{E}(Z\,|\,x) = (x)_{M_1} + \left(\frac{4x^4-6x^3+30x^2-20x+9}{12}\right)_{M_2} + \left(\frac{6x^4-30x^3+32x^2+9x}{12(2-x)}\right)_{N_1} + \left(\frac{14x^2-15x^3+6x^4-x^5}{3(2-x)}\right)_{N_2}$$

The variance is given by var $(Z\,|\,x) = \mathsf{E}(Z^2\,|\,x) - [\mathsf{E}(Z\,|\,x)]^2$.
The calculations when $k = 2$ in (4) are left for the reader.

Problems with Solutions Provide at the End of the Text

1*. Compute $E(X \mid Y = y)$ for the joint density

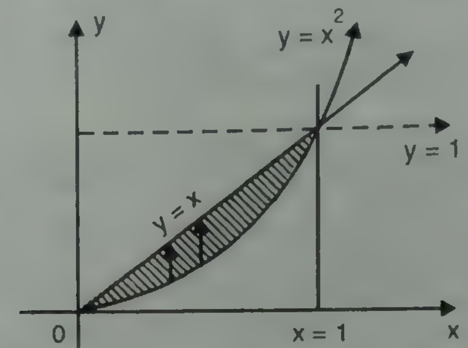
$$f(x, y) = y^{-1} e^{-y} e^{-x/y}, \quad 0 < x < \infty, \quad 0 < y < \infty.$$

2*. Give an example to show that $E(Y)$ may not exist though $E(XY)$ and $E(Y \mid x)$ may both exist.

3*. Show that $E[Y - g(X)]^2 \geq E[Y - E(Y \mid X)]^2$. ...(1)

4*. Find the expected number of random digits that should be generated to obtain three consecutive zeros.

5*. Variates X and Y are jointly distributed
$f(x, y) = 6$, if $x^2 \leq y \leq x, 0 < x < 1$, $f(x, y) = 0$, elsewhere. Find $E(Y \mid X = a)$ and $E(Y)$.



6*. X and Y are independent variates with integrating densities f_X, f_Y. Compute $P(X < Y)$ and find the distribution of $X + Y$.

7*. Two tetrahedra with faces numbered 1 to 4 are tossed. Let X denote the number on the downturned face of the first tetrahedron and Y the larger of the downturned numbers. Investigate the following :

(a) Joint density $f(x, y)$ of X, Y and marginals f_X and f_Y. Are X, Y independent ?

(b) Construct joint density $g(x, y)$ different from $f(x, y)$ but possessing same marginals f_X and f_Y.

(c) $F_{X,Y}(2, 3) = P\{X \leq 2, Y \leq 3\}$ (d) $\rho(X, Y)$ (e) $E(Y \mid X = 2)$.

8*. Given two variates X and Y with joint p.d.f. $f(x, y)$, prove that conditional mean of Y given X, coincides with its unconditional mean only if X and Y are independent. Show that the converse need not be true.

9*. The joint distribution of X and Y is as shown. Determine p and evaluate Cov (X, Y), Var (X) and Var (Y). Are X and Y independent ? Find the conditional p.d.f. of X given $Y = 0$ and compute Var $(Y \mid X = -1)$.

$y \rightarrow$ / $x \downarrow$	-1	0	1
-1	0	$2p$	0
0	p	$2p$	p
1	p	$2p$	p

10*. Variates X and Y have the joint distribution

$$f(x, y) = \begin{cases} e^{-y}(1-e^{-x}), & 0<x<y,\ 0<y<\infty \\ e^{-x}(1-e^{-y}), & 0<y<x,\ 0<x<\infty \end{cases} \qquad (x \rightleftharpoons y)$$

Verify that $f(x, y)$ is a bonafide p.d.f. and find the marginal distributions of X and Y. Also, find another joint p.d.f. having the same marginals. Evaluate :

(a) $E[Y \mid X = x]$, for $x > 0$. (b) $P(X \le 2, Y \le 2)$. (c) Corr (X, Y).

11*. Find the expectation for the square of the distance between two points selected at random on the boundary of a rectangle.

Exercises

1. (a) Let X and Y be variates of discrete type having p.m.f.

$f(x, y) = (x + 2y)/18$, $(x, y) = (1, 1), (1, 2), (2, 1), (2, 2)$; $f(x, y) = 0$, elsewhere.

Determine the conditional mean and variance of Y given $X = x$, when $x = 1$ or 2.

(b) Let $f(x, y) = (x + y)/32$, $x = 1, 2$, $y = 1, 2, 3, 4$. Find $E(X \mid Y)$. Are X, Y independent ?

[**Ans.** (a) $E(Y \mid 1) = 13/8$, $E(Y \mid 2) = 6/10$, $\text{Var}(Y \mid 1) = 15/64$, $\text{Var}(Y \mid 2) = 24/100$]

2. Variates X, Y have the joint p.m.f as :

(x, y)	$(0, 0)$	$(0, 1)$	$(1, 0)$	$(1, 1)$	$(2, 0)$	$(2, 1)$
$f(x, y)$	p	$3p$	$4p$	$3p$	$6p$	p

Determine p and find the two Marginal densities and the two Conditional means.

[**Ans.** $E(X \mid Y = 0) = 16/11$; $E(X \mid Y = 1) = 5/7$]

3. A joint p.m.f. of X and Y is given in the table.

(i) Find p, $E(X)$, $E(Y)$ and show that

$E(Y \mid X = 2) = 4/7$.

(ii) $P(X \le 2 \mid Y = 0) = 0.33$,

$P(Y \le 1 \mid X \le 2) = 17/20$.

$Y\downarrow$ \ $X\rightarrow$	0	1	2	3
0	$5p$	$8p$	$20p$	$10p$
1	$10p$	$15p$	$10p$	$5p$
2	$2p$	$5p$	$5p$	$5p$

4. Let $P(X = x, Y = y) = (x + 3y)/24$, where $(x, y) = (1, 1), (1, 2), (2, 1), (2, 2)$.

Show that the conditional mean and variance of X, given $Y = 2$ are $23/15$ and $56/225$.

5. Let $p(x, y) = \frac{1}{4}$, $(x, y) = (-3, -5), (-1, -1), (1, 1), (3, 5)$; $p(x, y) = 0$, elsewhere. Are X and Y are independent ? Show that $\rho_{XY} = 8/\sqrt{65}$ and $\text{Var}(X \mid y) = 20$.

6. Two variates have the following joint distribution, in which $2a + 2b + c = 1$. Are X, Y mutually independent ?

Calculate $E(X + Y)$, $\text{Var}(X + Y)$. Also find $\text{Var}(X \mid y)$ and $\text{Var}(Y \mid x)$.

$Y\downarrow$ \ $X\rightarrow$	-2	-1	0	1	2
0	0	0	c	0	0
1	0	b	0	b	0
2	a	0	0	0	0

[**Ans.** $2(2a + b)$, $4(b + 4a) - 4(b + 2a)^2$, $\text{Var}(X \mid y = 2) = 4$, $\text{Var}(Y \mid X = 1) = 0$]

7. Let $f(x, y) = 2, 0 < x < y < 1$; $f(x, y) = 0$, elsewhere.

   Show that $E(X \mid Y = y) = \frac{1}{2} y$, $\text{Var}(X \mid y) = y^2/12$.

8. The joint p.d.f. of (X, Y) is : $f(x, y) = a^2 e^{-a(x+y)}$, $x, y \geq 0$; $f(x, y) = 0$, elsewhere.

   Show that $E(X \mid X + Y \leq 4) = 4$.

9. The joint p.d.f. of (X, Y) is given to be

   $f(x, y) = k(4 - x - y)$, $0 \leq x, y \leq 2$, $f(x, y) = 0$, elsewhere.

   Find the constant k and the Marginal density functions of X and Y. Also find Conditional density functions and conditional means and show that

   $\text{Var}(X) = 11/3 = \text{Var}(Y)$ and $\text{Cov}(X, Y) = -1/36$. Verify $E(X \mid Y) = E(X)$.

   [**Ans.** $k = 1/8, f_1(t) = f_2(t) = (3 - t)/4, f_1(x \mid y) = (4 - x - y)/2(3 - y)$].

10. The joint p.d.f. of (X, Y) is $f(x, y) = 2 - x - y, 0 \leq x, y \leq 1, f(x, y) = 0$, elsewhere.

    Show that the Marginal densities are $f_1(t) = f_2(t) = [(3/2) - t] \, I(0 \leq t \leq 1)$.

    Obtain the Conditional densities and conditional means of X and Y and show that $\sigma_X^2 = 11/44 = \sigma_Y^2$ and $\sigma_{XY} = -1/144$.

11. The joint p.d.f. of (X, Y) is given to be

    $f(x, y) = 9(1 + x + y)/2(1 + x)^4 (1 + y)^4, 0 \leq x, y < \infty$, $f(x, y) = 0$, elsewhere.

    Find the Conditional densities and conditional means.

    [**Ans.** $E(Y \mid x) = (x + 3)/(3 + 2x)$, $E(X \mid y) = (y + 3)/(3 + 2y)$]

12. Let (X, Y) be jointly distributed with density

    (a) $f(x, y) = y(1 + x)^{-4} e^{-y(1+x)-1}$, $x, y \geq 0$; $f(x, y) = 0$, elsewhere.

    Find $E(Y \mid X)$, $\text{Corr}(X, Y)$, $F(x, y)$.

    (b) $f(x, y) = k(x + y)$, $0 < x < 1, 0 < y < 1$; $f(x, y) = 0$, elsewhere.

    Show that $k = 1$, $E(Y \mid x) = \left(\frac{1}{2}x + \frac{1}{3}\right) \Big/ \left(x + \frac{1}{2}\right)$, $\rho_{XY} = -\frac{1}{11}$, $F(x, y) = \frac{1}{2} xy(x + y)$.

13. The joint p.d.f. of X and Y is $f(x, y) = 3(x + y), 0 < x, y < 1, 0 \leq x + y < 1$.

    Show that : $f_1(t) = f_2(t) = 3(1 - t^2)/2$, $0 < t < 1$.

    (a) $P(X + Y < \frac{1}{2}) = 1/8$.

    (b) $E(Y \mid x) = (1 - x)(x + 2)/3(1 + x)$

    (c) $\text{Var}(Y \mid x) = (1 - x)^2 (x^2 + 4x + 1)/18(1 + x)^2$.

    (d) $\text{Corr}(X, Y) = 83/95$.

14. Joint p.d.f. of variates X and Y is given by

    $f(x, y) = \frac{1}{8}(6 - x - y)$, $0 < x < 2, 2 < y < 4$; $f(x, y) = 0$ elsewhere.

    Show that (a) $E(Y \mid x) = (52 - 18x)/6(3 - x)$ (b) $E(Y^2 \mid x) = (78 - 28x)/3(3 - x)$

    (c) $E(XY \mid X = x) = x(26 - 9x)/3(3 - x)$ (d) $\text{Corr}(X, Y) = -1/\sqrt{517}$. Also, verify that $E(Y) = E[E(Y \mid X)]$.

15. Are X and Y independent if X and Y have joint density :

    $$f(x, y) = 3/2, 0 < x < 1, -(x - 1)^2 < y < (x - 1)^2.$$

    Evaluate $E(X + Y)$, $E(X \mid Y = 0)$ and $E(Y \mid X = 1/2)$.

16. Joint density of X and Y is $f(x, y) = y^{-1} e^{-x/y} e^{-y}, 0 < x < \infty, 0 < y < \infty$. Show that $E(X \mid Y = y) = y$ and $E[X^2 \mid Y = y] = 2y^2$.

17. If $f(x, y) = e^{-y}/y, 0 < x < y$, $0 < y < \infty$, show that $E(X \mid Y) = y/2$ and $E(X^3 \mid y) = y^3/4$.

18. Let a point X be chosen from the interval $]0, 1[$ and then choose a point Y from the interval $]0, x[$, where x is the experimental value of X. With suitable assumptions about densities, compute $P(X + Y \geq 1)$ and $E(X \mid Y = y)$.

19. Let : $f(x, y) = 1/\pi a^2, x^2 + y^2 \leq a^2$; $f(x, y) = 0$, elsewhere.
Find the marginal distributions of X and Y and determine whether X and Y are independent ? Show that Cov $(X, Y) = 0$, $\text{Var}(Y \mid x) = (a^2 - x^2)/3$ and $\text{Var}(X \mid y) = (a^2 - y^2)/3$.

20. Let $X_1, X_2, ..., X_n$ be i.i.d. variates with common density.
$f(x) = (1/x^2)$. $1 < x < \infty$; $f(x) = 0$, elsewhere. Show that $E(X_1)$ does not exist.
Let $Y = \min \{X_1, X_2, ..., X_n\}$. Prove that $E(Y)$ exists and equals in $n/(n-1)$.

21. (a) If X and Y are continuous variates, show that $E[g(X, Y)] = E_Y\{E_{X \mid Y} g(X, Y)\}$.
(b) If X and Y are independent variates, then for any real-valued functions ϕ and ψ

$$E[\phi(X)\,\psi(Y)] = [E_X\,\phi(X)]\,[E_Y\,\psi(Y)]$$

(c) If V denotes the variance and $\psi = \psi\,(X, Y)$ is any function, then

$$\text{Var}(\psi) = E_y[V_{X \mid y}(\psi)] + V_y[E_{X \mid y}(\psi)].$$

22. (a) If $E(Y \mid X = x) = \mu$, where μ does not depend on x, show that $\text{Var}(Y) = E[\text{Var}(Y \mid X)]$.
(b) Prove or disprove :
If $E(Y \mid X) = X$, $E(X \mid Y) = Y$, $E(X^2) < \infty$, $E(Y^2) < \infty$, then $P(X = Y) = 1$.

23. Prove that Cov $(X, Y) =$ Cov $[X, E_{Y \mid X}(Y \mid X)]$. Also show that $\sigma_{XY} > 0$ (or < 0) according as $E(Y \mid X)$ is increasing (or decreasing) if all needed expectations exist.

24. Let $f(x, y) = ky^2, 0 < 2x < y < 2$; $f(x, y) = 0$, elsewhere. Find k, the conditional mean and variance of X given y.
Also, evaluate $P(2X + Y < 1)$ and $P(2X + Y < 1 \mid X > 1/2)$.

[**Ans.** $k = \frac{1}{2}$, $E(X \mid y) = y/4$, $\text{Var}(X \mid y) = y^2/48$, $p_1 = 29/1536$, $p_2 = 0$]

25. Let $f(x, y) = \frac{1}{2} e^{-y}, y > |x|, \quad -\infty < x < \infty$.
Show that $E(Y \mid x) = \exp(|x|)$ and $\text{Var}(Y \mid x) = 2 \exp(|x|) - \exp(2|x|)$.

26. Choose two real decimal numbers in the following way :
First a number N is chosen between 0 and 10. If $N < 4$, then a second number M is chosen uniformly between 0 and 3. If $N > 4$, then the second number M is chosen uniformly between 0 and 4. Let X and Y denote the random numbers chosen in the first and in the second case. Find $F(x, y), f(x, y)$, $E(Y \mid X = 1)$ and $E(X \mid Y = 2)$. Also compute
(a) $P(3 < X < 5 \mid 2 < Y < 4)$, (b) $P(2X + Y < 3)$, (c) $P(X > 2 \mid 2X + Y < 3)$.

Idealism is what precedes experience; cynicism is what follows.

(David T. Wolf)

There is nothing as deceptive as an obvious fact. (Canon Doyle)

Generating Functions (or Transform Methods)

8

8-00. *Generating Functions*

For a given collection of numbers $\langle a_k, k \geq 0 \rangle$, we define

$$\text{(i) } G_a(t) = \sum_{k=0}^{\infty} a_k t^k; \qquad \text{(ii) } H_a(t) = \sum_{k=0}^{\infty} \frac{a_k t^k}{k!}$$

which are called (i) generating function of $\langle a_k \rangle$ and (ii) exponential generating function of $\langle a_k \rangle$. These were introduced by Euler and De Moivre in the early 18th Century.

Uniqueness Theorem (Crucial result)

If $G_a(t) = G_b(t)$ for $x_0 < t < x_1$, then $\langle a_n \rangle = \langle b_n \rangle, \forall\ n.$

Convolution Theorem **:** If $a_n = \sum_{j=0}^{n} c_j b_{n-j},\ n \geq 0$, then

$$G_a(t) = G_c(t) \,.\, G_b(t)$$

8-10. *Moment Generating Function*

Definition. The moment generating function (written m.g.f.) of a random variable X is denoted by M_x and is defined by the relation

$$\underline{M_x(t) \equiv M(t : X) = \mathsf{E}(e^{tX})} \qquad \text{...(1)}$$

provided this expected value exists for every value of t in some interval $-h < t < h$; $h > 0$. According as X is a discrete or a continuous variate, the Eq. (1) is written as

$$\underline{M(t:X) = \Sigma_x f(x)\, e^{tx}, \text{ or } M(t:X) = \int_{-\infty}^{\infty} f(x)\, e^{tx}\, dx} \qquad \text{(Kernel Form) ...(2)}$$

where $f(x)$ is the discrete or continuous density funciton of X. Expression (2) is sometimes called Dirichlet's form (series).

If a m.g.f. exists, then $M_X(t)$ is continuously differentiable in some neighbourhood of the origin. Differentiating (2) n times w.r.t. 't', we get

$$\frac{d^n}{dt^n} M_X(t) = \sum_x f(x)\, e^{tx} \,.\, x^n \Rightarrow M^n(0) = \left[\frac{d^n}{dt^n} M_X(t)\right]_{t=0} = \mathsf{E}(X^n) = \mu_n'$$

From Eq. (3) it follows that the moments of a distribution may be obtained from the m.g.f. function by differentiation and this justifies the reason for the name : m.g.f.

Note **:** The function M_X, called m.g.f. of X, assumes for the real variable t, the value $M_X(t)$ or $M(t : X)$. The notation $M(t : X)$ is neater when m.g.f. for combinations like $aX + bY + c$, etc. are used.

8-11. *Taylor's Series of M.G.F. about an Arbitrary Point*

The m.g.f. of a r.v. $Y = (X - a)/b$ is defined by

$$M(t : Y) = \mathsf{E}[e^{t(X-a)/b}] \left[M_Y(t) = \int_{-\infty}^{\infty} f(x)\, e^{t(x-a)/b} dx \text{ or } \sum_x f(x)\, e^{t(x-a)/b} \right] \quad \ldots(1)$$

Differentiating (1) w.r.to. 't' r times, assuming interchange of differentiation and Expectation, gives

$$\frac{d^r}{dt^r} M(t:Y) = \frac{d^r}{dt^r} \mathsf{E}[e^{t(X-a)/b}] = \mathsf{E}\left\{ \frac{d^r}{dt^r} e^{t(X-a)/b} \right\} = \mathsf{E}\left\{ \left(\frac{X-a}{b} \right)^r . e^{t(x-a)/b} \right\}$$

Put $t = 0$, on either side to get

$$D^r M_Y(0) = \mathsf{E}\left[\frac{(X-a)^r}{b^r} \right] = \frac{\mathsf{E}(X-a)^r}{b^r} = \frac{\mu_r'}{b^r} \qquad [\mu_r' = \mathsf{E}(x-a)^r] \quad \ldots(2)$$

Now Taylor's series for any function $M_Y(t)$ is

$$M_Y(t) = M_Y(0) + M_Y'(0)\, t + M_Y''(0) \frac{t^2}{2!} + \ldots + M_Y^{(r)}(0) \frac{t^r}{r!} + \ldots = \sum_{r=0}^{\infty} M_Y^{(r)}(0) . \frac{t^r}{r!} \quad \ldots(3)$$

Substituting from (2) into (3), we get

$$M_Y(t) = \sum_{r=0}^{\infty} \frac{\mu_r'}{r!} \left(\frac{t}{b} \right)^r = \sum_{r=0}^{\infty} \frac{\mu_r'}{b^r} . \frac{t^r}{r!} \quad \ldots(4)$$

Special Taylor Series :

(i) $M_X(t) = \sum_{r=0}^{\infty} \frac{\mu_r'}{r!} t^r$ [MGF of a r.v. X, ($a = 0$, $b = 1$) $\mu'_r = \mathsf{E}(X^r)$]

(ii) $M_\mu(t) = \sum_{r=0}^{\infty} \frac{\mu_r}{r!} t^r$ [MGF of centered r.v. $X - \mu$, ($a = \mu$, $b = 1$) $\mu_r = \mathsf{E}(X - \mu)^r$]

(iii) $m(t) = \sum_{r=0}^{\infty} \frac{\mu_r}{r!} \frac{t^n}{\sigma^r}$ [MGF of standard r.v. $(X - \mu)/\sigma$, ($a = \mu$, $b = \sigma$) $\mu_r = \mathsf{E}(X - \mu)^r$]

The series or integral in (1) may not exist, (i.e. not converge to a finite value) for all values of t, hence $M_X(t)$ is not defined for all values of t. However, at $t = 0$, m.g.f. *always exists* and equals 1.

Comments. It is possible that a variate X may neither possess moments nor m.g.f., e.g. consider a density : $f(x) = 1/x^2$, $1 \le x < \infty$; $f(x) = 0$, elsewhere.

Here $M(t:X) = \int_1^{\infty} (e^{tx} / x^2)\, dx.$

For $t > 0$, this integral is divergent and hence $M_X(t)$ does not exist. Simple moments μ_r', $r \ge 1$, also do not exist.

Example : Show that a m.g.f. may exist without the existence of moments.

Solution. *Yule's Distribution.* Let $f(x) = 1/x(x + 1)$, $x = 1, 2, 3, \ldots,$; $f(x) = 0$, elsewhere. [Let us write $T = e^t$]

$$M(t:X) = \sum_{x=1}^{\infty} \frac{e^{tx}}{x(x+1)} = \sum_{x=1}^{\infty} T^x \left(\frac{1}{x} - \frac{1}{x+1}\right) = \left(\sum_{x=1}^{\infty} \frac{T^x}{x}\right) - \frac{1}{T}\left(\sum_{x=1}^{\infty} \frac{T^{x+1}}{x+1}\right)$$

$$= \sum_{x=1}^{\infty} \frac{T^x}{x} - \frac{1}{T}\left[\sum_{x=1}^{\infty} \frac{T^x}{x} - T\right] = 1 + (1 - T^{-1}) \sum_{x=1}^{\infty} \frac{T^x}{x}$$

$$= 1 + (1 - T^{-1})[-\log(1-T)] = 1 - (1 - e^{-t}) \log(1 - e^t).$$

If $t = 0$, $M(0) \to 1$; if $t < 0$, $|e^t| < 1$; so $M(t:X)$ exists for $t < 0$; if $t > 0$, $\log(1 - e^t)$ is non-existent and $M(t:X)$ does not exist. Thus mgf $M(t)$ does exist for some values of t (viz. for $t \le 0$) but not every t. Since

$$\mathrm{E}(X) = \sum_{x=1}^{\infty} \frac{x}{x(1+x)} = \sum_{x=1}^{\infty} \frac{1}{x+1} \to \infty, \quad \text{[Divergent Harmonic series]}$$

it follows that μ'_1 and hence all higher moments, do not exist.

8-12. *Effect of Linear Transformations on Simple m.g.f.*

$M_{a+bx}(t)$

$$\boxed{M(t:a+bX) = e^{at}\, M(bt:X)} \qquad (a,\ b \text{ constants}) \qquad \ldots(1)$$

Proof. $M(t:a+bX) = \mathrm{E}[e^{t(a+bX)}] = \mathrm{E}[e^{at} \,.\, e^{(bt)X}] = e^{at}\mathrm{E}[e^{(bt)X}] = e^{at} M(bt:X)$. [By Lin E]

Cor. For a *standardized* variate $X^* = (X - \mu)/\sigma$, we have

$$M(t:X^*) = M[t:(X/\sigma) - (\mu/\sigma)] = e^{-\mu t/\sigma} M[(t/\sigma):X]. \qquad \ldots(2)$$

Note. The *effect of change of origin and scale* on m.g.f. is governed by (1).

8-13. *Linear Combination Property*

If X and Y are two *independent variates and a, b, c, constants, then*

$$\boxed{M(t:aX+bY+c) = e^{ct} M(at:X)\, M(bt:Y)} \qquad \ldots(1)$$

Proof. $M(t : aX + bY + c) = \mathrm{E}[e^{t(aX+bY+c)}] = \mathrm{E}[e^{atX} \,.\, e^{btY} e^{ct}]$

$$= e^{ct} \,.\, \mathrm{E}[e^{atX}] \,.\, \mathrm{E}[e^{btY}] \qquad [\varphi(X) \text{ and } \psi(Y) \text{ are independent}]$$

$$= e^{ct} \,.\, M(at:X)\, M(bt:Y) \qquad \text{[by def. of m.g.f.]}$$

In particular, the formulae for the sum of two indep. variates or a constant multiple are

$$M(t:X+Y) = M(t:X) \,.\, M(t:Y); \quad M(t:aX) = M(at:X).$$

Generalization : If X_i are all independent variates, then

$$M(t:a_1X_1 + a_2X_2 + \ldots + a_nX_n) = M(a_1t:X_1) \,.\, M(a_2t:X_2) \ldots M(a_nt:X_n)$$

The proof is immediate by mathematical induction.

8-14. *Uniqueness Theorem*

The m.f.g. uniquely determines a c.d.f., and conversely if the m.f.g. exists, it is unique.

Note. **Differentiation under expectation sign (DUES)**

If $M(t:X)$ exists for $-h < t < h$; $h > 0$, the derivatives of all orders exist at $t = 0$, and can be evaluated under the expectation sign. Thus

$$[M^{(r)}(t)]_{t=0} \equiv M^{(r)}(0) = \mathsf{E}(X^r), \qquad r \text{ positive integer.}$$

Proofs of these results are rather involved, and are omitted.

Note : If m.g.f. does not exist for all $t \in]-h, h[$, $h > 0$, then $M(t)$ need not generate moments. See Example 1 (Yule distribution).

Example : Show that existence of moments, all or some, need not imply the existence of m.f.g.

Solution. (i) Consider the Pareto distribution : $F(x) = 1 - (a/x)^{\theta}, x \geq a, \theta > 1.$

$$\mathsf{E}(X^r) = \int_a^{\infty} \frac{a^{\theta}\theta}{x^{\theta+1}} x^r \, dx = \theta \, a^{\theta} \int_a^{\infty} \frac{1}{x^{(\theta - r)+1}} dx.$$

This integral is convergent iff $r < \theta$, and then $\mathsf{E}(X^r) = \theta a^r / (\theta - r), \theta > r$.

However : $M(t:X) = \int_a^{\infty} \frac{a^{\theta}\theta}{x^{\theta+1}} e^{tx} \, dx = \theta a^{\theta} \int_a^{\infty} e^{tx} \, x^{-(\theta+1)} \, dx \to \infty$

becuase $(e^{tx} / x^{\theta+1}) > x^k$, for some positive integer k. Thus, m.f.g. does not exist.

Note : (i) Student's t-distribution and Fisher's F-distribution have moments of all orders, (in the domain of their definitions) but their m.g.f.'s do not exist.

(ii) Let X be a r.v. with $P(X = \pm 2^n)^{2k} = 1/n\,!\,(2e)$; $n = 0, 1, 2, \ldots$ Then all moments of X exist, viz.

$$\mathsf{E}(X^{2k+1}) = 0,\ \mathsf{E}(X^{2k}) = \sum_{n=0}^{\infty} (\pm 2^n)^{2k} \cdot \frac{(e)^{-1}}{n!} = \left(\sum_{n=0}^{\infty} \frac{(2^{2k})^n}{n!} \right) e^{-1} = e^{-1} \cdot e^{2k} = \exp(2^{2k} - 1).$$

$$M(t) = \sum_{n=0}^{\infty} e^{\pm 2^n t} P(Y = \pm 2^n) = \sum_{n=0}^{\infty} \left(\frac{e^{2^n t} + e^{-2^n t}}{2} \right) \frac{e^{-1}}{n!} = e^{-1} \sum_{n=0}^{\infty} \frac{\cosh(2^n t)}{n!}. \qquad \ldots(1)$$

If m.g.f. of X exists, it is given by (1). However, the series in (1) converges only for $t = 0$, which we can verify by using D'Alemberts' ratio test.

Hence, $M(t)$ is defined only for $t = 0$ and thus existence of all the moments of r.v. X does not imply that m.g.f. exists.

Properties of M.G.F. : $M_X(t) = \mathsf{E}(e^{tX})$

1. $M_X(0) = 1,\ M(t) > 0, \forall\, t$ where $M(t)$ is defined
2. For $X \geq 0$, $M'(t) \geq 0$ and $M_X(t) \uparrow$, For $X \leq 0$, $M'(t) \leq 0$ and $M_X(t) \downarrow$
3. $M_X^{(k)}(0) = \mathsf{E}(X^k) \equiv \mu_k',\ k = 0, 1, 2, 3, \ldots$
4. $M_X(t) = \Sigma(t^r / r!)\, M^{(r)}(0) = \Sigma(t^r / r!)\, \mu_r' = 1 + \mu_1' t + \mu_2' (t^2/2!) + \ldots$ Taylor series
5. $M(t : aX + b) = e^{bt} \cdot M(at : X)$
6. $M(t : X + Y) = M(t : X) \cdot M(t : Y)$ if X and Y are independent.
7. $M_X(t)$ characterizes the distribution of X under reasonable conditions.

8-16. *M.G.F.s of Some Standard Distribitions*

Some frequently occuring probability laws are stated in §3-40 and §3-60. The expressions for m.g.f.s corresponding to some of these distributions are obtained in this section. These results are extensively used in the sequel.

1. ***Degenerate distribution.*** If $P(X = c) = 1$, $P(X \neq c) = 0$, then

$$M(t:X) = \mathsf{E}(e^{tX}) = 1 \,.\, e^{ct} + 0 \,.\, e^{c't} = e^{ct}.$$

2. ***Bernoulli's distribution.*** $f(x) = q^{1-x} p^x, \quad x = 0, 1, \quad p + q = 1$

$$M(t) = \mathsf{E}(e^{tX}) = q \,.\, e^{0t} + p . e^t = (q + pe^t), \; [X \sim \text{Ber}(p)]$$

2′. ***Binomial distribution.*** $X \sim \text{bin}\,(n, p); \quad f(x) = \binom{n}{x} q^{n-x} p^x, \quad x = 0, 1, 2, \ldots, n.$

$$M(t:X) = \mathsf{E}(e^{tX}) = \sum_{x=0}^{n} \binom{n}{x} q^{n-x} p^x e^{tx} = \sum_{x=0}^{\infty} \binom{n}{x} q^{n-x} (pe^t)^x = (q + pe^t)^n. \text{ [by Bin. expansion]}$$

Note. bin $(1, p)$ = Ber(p).

3. ***Poisson distribution.*** $X \sim \text{Pois}\,(\lambda); \quad f(x) = e^{-\lambda} \lambda^x / x!, \; \lambda > 0, \quad x = 0, 1, 2, \ldots .$

$$M(t:X) = \mathsf{E}(e^{tX}) = \sum_{x=0}^{n} \frac{e^{-\lambda}\lambda^x}{x!} e^{tx} = e^{-\lambda} \sum_{x=0}^{n} \frac{(\lambda e^t)^x}{x!} = e^{-\lambda} e^{\lambda e^t} = e^{\lambda(e^t - 1)}. \text{ [by exponential series]}$$

4. ***Geometric distribution.*** $X \sim$ geom (p); $\quad f(x) = q^x p, \qquad x = 0, 1, 2, \ldots$

$$M(t:X) = \mathsf{E}(e^{tX}) = \sum_{x=0}^{\infty} (q^x p)\, e^{tx} = p \sum_{x=0}^{\infty} (qe^t)^x = \frac{p}{1 - qe^t}; \qquad |\, qe^t \,| < 1. \qquad \text{[By G.P.]}$$

4′. ***Pascal Geometric distribution.*** $X \sim \text{gem}\,(p)\,; f(x) = q^{x-1} p, \qquad x = 1, 2, 3, \ldots$

$$M(t:X) = \mathsf{E}(e^{tX}) = \sum_{x=0}^{\infty} (q^{x-1} p)\, e^{tx} = pe^t \sum_{x=0}^{\infty} (qe^t)^{x-1} = \frac{pe^t}{1 - qe^t}, \qquad |\, qe^t \,| < 1. \quad \text{[By G.P.]}$$

5. ***Neg-bin distribution.*** $X \sim NB(k, p); \; f(x) = \binom{x + k - 1}{k - 1} p^k q^x \equiv \binom{-k}{x} p^k (-q)^x, \quad x = 0, 1, 2, \ldots$

$$M_X(t) = \mathsf{E}(e^{tX}) = \sum_{x=0}^{\infty} \binom{x + k - 1}{k - 1} q^x p^k e^{tX} = p^k \sum_{x=0}^{\infty} \binom{x + k - 1}{x} (qe^t)^x = p^k (1 - qe^t)^{-k}.$$

$$= [p / (1 - qe^t)]^k.$$

5′. ***Pascal N-B distribution.*** $Y \sim NB^*(k, p); \quad f(y) = \binom{y - 1}{k - 1} p^k q^{y-k}, \quad y = k, k + 1, k + 2, \ldots$

$$M(t:Y) = \mathsf{E}(e^{tY}) = \sum_{y=k}^{\infty} \binom{y-1}{k-1} p^k q^{y-k} \,.\, e^{ty} = (pe^t)^k \sum_{x=0}^{\infty} \binom{k + x - 1}{x} (qe^t)^x = (pe^t)^k (1 - qe^t)^{-k} = \left(\frac{pe^t}{1 - qe^t} \right)^k.$$

We have used $y = k + x$. Here $Y = k + X$, i.e. $NB^* = k + NB$.

6. ***Multinomial distribution.***

$$X \sim \text{Mul}\,(n, k; p_i), \;\; f(x_1, \ldots, x_k) = \binom{n}{x_1, \ldots, x_k} p_1^{x_1} \ldots p_k^{x_k}, 0 < p_i < 1$$

$$M(t:X) = \mathsf{E}(e^{t_1 x_1 + \ldots + t_k x_k}) = \sum_{x_i} \binom{n}{x_1, \ldots, x_k} p_1^{x_1} \ldots p_k^{x_k} \,.\, e^{t_1 x_1 + \ldots + t_k x_k}$$

$$= \sum_{x_i} \binom{n}{x_1, \ldots, x_k} (p_1 e^{t_1})^{x_1} \ldots (p_k e^{tk})^{xk} = (p_1 e^{t_1} + p_2 e^{t_2} + \ldots + p_k e^{t_k})^n.$$

7. ***Uniform distribution.*** $X \sim$ Unif (a, b); $f(x) = 1/(b-a)$, $\quad a \le x \le b$

$$M(t : X) = \mathrm{E}(e^{tX}) = \int_a^b \frac{e^{tx}}{b-a} dx = \frac{e^{bt} - e^{at}}{(b-a)t} \qquad (t \neq 0).$$

Note. For $U(0, 1)$, $M(t : X) = (e^t - 1)/t$, $t \neq 0$. [Standard uniform distribution]

Note. Frequently we write U = unif (0, 1)

8. ***Cumulative distribution function*** (c.d.f.) $Y = F(X)$. X continuous.

$$M(t : F(X)) = \mathrm{E}(e^{tF(X)}) = \int_{-\infty}^{\infty} e^{tF(x)} F'(x)\, dx = \int_0^1 e^{ty} dy = \frac{(e^t - 1)}{t} \quad \text{[mgf of unif(0, 1)]}$$

Note that $F(\infty) = 1$, $F(-\infty) = 0$, $F'(x) = f(x)$.

9. ***Normal (or Gaussion) distribution.*** $X \sim N(\mu, \sigma^2)$

$$f(x) = (\sigma\sqrt{2\pi})^{-1} e^{-(x-\mu)^2/2\sigma^2}, \qquad -\infty < x < \infty.$$

$$M(t : X) = \mathrm{E}(e^{tX}) = \int_{-\infty}^{\infty} \frac{e^{-(x-\mu)^2/2\sigma^2}}{\sqrt{2\pi}\sigma} . e^{tx}\, dx = \int_{-\infty}^{\infty} \frac{e^{-z^2/2} . e^{t(\mu + \sigma z)}}{\sqrt{2\pi}} dz. \qquad \left[z = \frac{x-\mu}{\sigma}\right]$$

Since : $-\frac{1}{2}z^2 + t(\mu + \sigma z) = (\mu t + \frac{1}{2}\sigma^2 t^2) - \frac{1}{2}(z - \sigma t)^2 = (\mu t + \frac{1}{2}\sigma^2 t^2) - \frac{1}{2}u^2$, $\qquad [u = z - \sigma t]$

$$\therefore \quad M(t : X) = e^{\mu t + \frac{1}{2}\sigma^2 t^2} \left(\int_{-\infty}^{\infty} \frac{e^{u^2/2} du}{\sqrt{2\pi}} \right) = e^{\mu t + \frac{1}{2}\sigma^2 t^2}. \qquad \text{[Area under } N(0, 1) \text{ is unity]}$$

Standard Normal distribution. $X \sim N(0, 1)$. $M(t : X) = e^{t^2/2}$

10. ***Gamma Distribution.*** $X \sim$ gam (α, λ). $f(x) = \lambda^\alpha e^{-\lambda x} x^{\alpha - 1} / \Gamma(\alpha)$, $\quad x > 0, \alpha > 0$.

$$M(t : X) = \mathrm{E}(e^{tX}) = \int_0^{\infty} \frac{\lambda^\alpha e^{-\lambda x} x^{\alpha-1}}{\Gamma(\alpha)} e^{tx}\, dx = \frac{\lambda^a}{\Gamma(\alpha)} \int_0^{\infty} e^{-(\lambda - t)x} x^{\alpha - 1} dx = \frac{\lambda^\alpha}{\Gamma(\alpha)} \frac{\Gamma(\alpha)}{(\lambda - t)^\alpha} = \left(1 - \frac{t}{\lambda}\right)^{-\alpha}.$$

Note. gam $(1, \lambda) \equiv$ expo (λ), gam $(\frac{1}{2}n, \frac{1}{2}) \equiv \chi^2_{(n)}$.

11. ***Exponential distribution.*** $X \sim$ Expo (λ). $f(x) = \lambda e^{-\lambda x}, \lambda > 0, x > 0$

$$M(t : X) = \mathrm{E}(e^{tX}) = \int_0^{\infty} \lambda e^{-\lambda x} . e^{tx} dx = \lambda \int_0^{\infty} e^{-(\lambda - t)x} dx = \frac{\lambda}{\lambda - t} = \left(1 - \frac{t}{\lambda}\right)^{-1}.$$

Note $(-1/\lambda) \ln U \sim$ expo (λ); $\quad X \sim$ expo (λ), $cX \sim$ expo (λ / c), $c > 0$

12. ***Chi-square distribution.*** $X \sim \chi^2_{(n)} \cdot f(x) = \dfrac{e^{-x/2} . x^{(n/2)-1}}{2^{n/2}\Gamma(n/2)}, \qquad x > 0.$

$$M(t : X) = \mathrm{E}(e^{tX}) = \int_0^{\infty} \frac{e^{-x/2} . x^{(n/2)-1}}{2^{n/2} P(n/2)} . e^{tx} = \int_0^{\infty} \frac{e^{-1/2(1-2t)x} x^{(n/2)-1}}{2^{n/2} P(n/2)} dx = \frac{\Gamma(n/2)}{2^{n/2}\Gamma(n/2)[(1-2t)/2]^{n/2}} = (1-2t)^{-n/2}$$

13. ***Laplace distribution.*** $X \sim$ Lap (μ, λ). $f(x) = \frac{1}{2}\lambda e^{-\lambda|x-\mu|}, -\infty < x < \infty, \lambda > 0.$

$$M(t : X) = \mathrm{E}(e^{tX}) = \frac{1}{2}\lambda \int_{-\infty}^{\infty} e^{-\lambda|x-\mu|} e^{tx}\, dx = \frac{1}{2}\lambda e^{\mu t} \int_{-\infty}^{\infty} e^{-\lambda|z|} e^{tz}\, dz \qquad [z = x - \mu]$$

$$= \frac{1}{2}\lambda e^{\mu t}\left\{\int_{-\infty}^{0} e^{(\lambda+t)z} dz + \int_{0}^{\infty} e^{-(\lambda-t)z} dz\right\} = \frac{1}{2}\lambda e^{\mu t}\left\{\int_{0}^{\infty} e^{-(\lambda+t)y} dy + \int_{0}^{\infty} e^{-(\lambda-t)z} dz\right\}$$

$$= \frac{1}{2}\lambda e^{\mu t}\{1/(\lambda+t)\} + (1/(\lambda-t))\} = \lambda^2 e^{\mu t}/(\lambda^2 - t^2) = e^{\mu t}(1 - t^2/\lambda^2)^{-1} = e^{\mu t}\lambda^2/(\lambda^2 - t^2)$$

8-17. *Laplace Transform*

Let the r.v. X be non-negative. The *Laplace Transform* of X is denoted by Lap $(t : X)$ and is defined by

$$\text{Lap}\,(t : X) = \mathrm{E}(e^{-tX}) = M(-t : X) \quad t > 0, X \geq 0.$$

We shall sometimes evaluate $M(-t : X)$. An extensive literature and Tables of Laplace Transforms (L.Ts) are readily available.

Illustration. $X \sim$ Expo (a) and $Y \sim$ Expo (b) be independently distributed ($a \neq b$). To find the distribution of $Z = X + Y$, we use Lap-transform (L.T.).

$$M(-t : Z) = M(-t : X + Y) = M(-t : X).M(-t : Y) = \left(1 + \frac{t}{a}\right)^{-1}\left(1 + \frac{t}{b}\right)^{-1} = \frac{ab}{b-a}\left(\frac{1}{t+a} - \frac{1}{t+b}\right). \quad \ldots(1)$$

$$\text{Let } f(z) = \frac{ab}{b-a}(e^{-az} - b^{-bz}),\ a \neq b \quad 0 < z < \infty. \quad \ldots(2)$$

We trially find that Lap-Trans of distribution (2) is given by (1). Hence (2) is the density of $X + Y$.

8-18. *Sums of Independent Random Variables : Linear Combinations*

1. Let $X \sim N(\mu_1, \sigma_1^2)$ and $Y \sim N(\mu_2, \sigma_2^2)$ be independent. Then

$$(X + Y) \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2),$$

Linear Combination. Let $X \sim N(\mu_1, \sigma_1^2)$ and $Y \sim N(\mu_2, \sigma_2^2)$ be independent. Then

$$\boxed{(aX + bY + c) \sim N(a\mu_1 + b\mu_2 + c, a^2\sigma_1^2 + b^2\sigma_2^2).}$$

2. Let $X \sim$ gam (a, λ) and $Y \sim$ gam (b, λ) be independent. Then $(X + Y) \sim$ gam $(a + b, \lambda)$.

3. Let $X \sim$ Expo (λ) and $Y \sim$ Expo (λ) be independent. Then $(X + Y) \sim$ gam $(2, \lambda)$.

Proof. In each case, we use $M(t : X + Y) = M(t : X)\,M(t : Y)$ and appeal to uniqueness theorem.

1. $M_X(t) = e^{\mu_1 t + (\sigma_1^2 t^2/2)}, M_Y(t) = e^{\mu_2 t + (\sigma_2^2 t^2/2)}, M_X(t).M_Y(t) = e^{(\mu_1 + \mu_2)t + (\sigma_1^2 + \sigma_2^2)t^2/2} = M(t : X + Y)$

This shows that $(X + Y) \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$

Linear Combination. $M(t : aX + bY + c) = e^{ct} M(at : X)\,M(bt : Y) = e^{ct} . e^{\mu_1 at + (a^2\sigma_1^2 t^2/2)} . e^{\mu_2 bt + (b^2\sigma_2^2 t^2/2)}$

$$= \exp\,[a\mu_1 + b\mu_2 + c + (a^2\sigma_1^2 + b^2\sigma_2^2)/2]$$

Thus linear combination of Normal variates are again Normal

2. $M_X(t) = (1 - t/\lambda)^{-a}, M_Y(t) = (1 - t/\lambda)^{-b}, M_X(t).M_Y(t) = (1 - t/\lambda)^{-(a+b)} = M(t : X + Y).$

This shows that $(X + Y) \sim$ gam $(a + b, \lambda)$

3. $M_X(t) = (1 - t/\lambda)^{-1} = M_Y(t), M_X(t).M_Y(t) = (1 - t/\lambda)^{-2} = M(t : X + Y).$

This shows that $(X + Y) \sim \text{gam}(\alpha, \lambda)$. Sum of Expo variates is not an Expo variate.

Note. Sums of indep. binomial, Poisson and geometric variates are considered in §8-54 & 14.84

Sum of two Cauchy variates, using Ch. functions, is considered in §9-71

Sum of several uniformly distributed variates is considered in §18-30

8-19. *Worked-out Problems*

Example 1. If $M(t:X) = \frac{2}{5}e^t + \frac{1}{5}e^{2t} + \frac{2}{5}e^{3t}$, find Var (X) and p.m.f. of the variate X. Compute μ'_r as well.

Solution. The m.g.f. of variate X is given by Dirichlet's form

$$M(t) = \mathsf{E}(e^{tX}) = \Sigma f(k)\, e^{tk},\ k = 1, 2, 3, \ldots$$

Comparing this with the given $M(t)$, equating components, we get p.m.f.

$$P(X = 1) = 2/5,\ P(X = 2) = 1/5,\quad P(X = 3) = 2/5,\quad P(X \geq 4) = 0.$$

We now use Art. 8-10 (3) to find Var (X).

$$M'(t) = (2/5)\,e^t + (2/5)e^{2t} + (6/5)e^{3t},\quad M'(0) = \mathsf{E}(X) = 2.$$

$$M''(t) = (2/5)e^t + (4/5)e^{2t} + (18/5)e^{3t},\quad M''(0) = \mathsf{E}(X^2) = 24/5.$$

$$\therefore\quad \text{Var}\,(X) = \mathsf{E}(X^2) - \mathsf{E}^2(X) = M''(0) - [M'(0)]^2 = (24/5) - 4 = 4/25.$$

Note. Using the exponential expansions and collecting coefficients provide

$$M(t) = \tfrac{1}{5}\sum_{r=0}^{\infty}\frac{t^r}{r!}(2 + 2^r + 2(3)^r] \Rightarrow \mu'_r = \tfrac{1}{5}[2 + 2^r + 2(3)^r].$$

Thus $\mu'_1 = 2,\ \mu'_2 = 24/5,\ \sigma_X^2 = (24/5) - 4 = 4/25.$

Example 2. Find the m.g.f. of logastics distribution and hence obtain its mean and variance.

Solution. Here $F(x) = \{1 + \exp[-(x - \mu)/k]\}^{-1}$. [Logastics c.d.f.]

$$f(x) = \exp[-(x - \mu)/k]/k\{1 + \exp[-(x - \mu)/k]\}^2,\quad -\infty < n < \infty \qquad \ldots(1)$$

$$\therefore\quad M_X(t) = \mathsf{E}(e^{tX}) = \int_{-\infty}^{\infty} e^{tX} f(x)\,dx \qquad \ldots(2)$$

Put $e^{-(x-\mu)/k} = u$, then $e^{-(x-\mu)k}dx = -k\,du,\quad e^x = e^{\mu}u^{-k}$.

Substituting into (2), we get

$$M(t) = \int_0^{\infty}\frac{e^{\mu t}\,.\mu^{-kt}du}{(1+u)^2} \qquad \left[\int_0^{\infty}\frac{y^{a-1}dy}{(1+y)^{a+b}} = B(a, b)\right]$$

$$= e^{\mu t}\,B(1 - kt,\ 1 + kt)$$

$$= e^{\mu t}\Gamma(1 - kt)\,\Gamma(1 + kt), \qquad [\Gamma(p)\,\Gamma(1 - p) = \pi/\sin p\pi, 0 < p < 1]$$

$$= (kt)\,e^{\mu t}\Gamma(p)\,\Gamma(1 - p) \qquad [p = 1 - kt, \Gamma(2 - p) = kt\Gamma(1 - p)]$$

$$= e^{\mu t}(\pi kt)/\sin(\pi kt)$$

$$K(t) = \ln M(t) = \mu t - \ln[\sin(\pi kt)/(\pi kt)] \qquad [\text{Expand } \sin\theta]$$

$$= \mu t - \ln[1 - \pi^2 k^2 t^2 / 3! \ldots] \qquad \text{[Use } \ln\text{-series]}$$

$$= \mu t + \pi^2 k^2 t^2 / 6 + \ldots$$

$$\mathrm{E}(X) = k_1 = \mu, \qquad \mathrm{Var}(X) = k_2 = \pi^2 k^2 / 3, \ldots$$

***Example* 3.** Find the m.g.f. of the *Inverse Gaussian* density :

$$f(x) = \left(\frac{\lambda}{2\pi x^3}\right)^{1/2} e^{-\lambda(x-\mu)^2/2x\mu^2}, x > 0, \mu > 0, \lambda > 0.$$

Solution.

$$M_X(t) = \left(\frac{\lambda}{2\pi}\right)^{1/2} \int_0^\infty e^{tx} e^{-\lambda(x-\mu)^2/2x\mu^2} \frac{dx}{x^{3/2}} \qquad [\text{Put } x = y^{-2}]$$

$$= 2e^{\lambda/\mu}\left(\frac{\lambda}{2\pi}\right)^{1/2} \int_0^\infty e^{-(\lambda y^2/2) - [(\lambda/2\mu^2) - t]/y^2} dy$$

Since : $\int_0^\infty e^{-a^2x^2 - (b^2/x^2)} dx = \frac{\sqrt{\pi}}{2a} e^{-2ab}, a > 0, b \geq 0$

$$\therefore \qquad M(t) = \exp\left\{\frac{\lambda}{\mu}\left[1 - \left(1 - \frac{2\mu^2 t}{\lambda}\right)^{1/2}\right]\right\}. \qquad \left[\text{use } a^2 = \frac{\lambda}{2}, \ b^2 = \frac{\lambda}{2\mu^2} - t\right]$$

Note. We can find the mean and variance using cumulants

$$K(t) = \ln M(t) = \frac{\lambda}{\mu}\left\{1 - \left(1 - \frac{2\mu^2 t}{\lambda}\right)^{1/2}\right\} = \mu t + \frac{\mu^3}{\lambda}\frac{t^2}{2!} + \ldots$$

Thus $k_1 = \mathrm{E}(X) = \mu$, $k_2 = \mathrm{Var}(X) = \mu^3/\lambda$

Remark. The variate $Z = \lambda(X - \mu)^2 / X\mu^2 \sim \chi^2_{(1)}$ ...(1)

Problems with Solutions Provided at the End of the Text

1*. Explain why there can be no r.v. X for which $M(t : X) = t/(1 - t)$.

2*. Show that for the rectangular dist. : $f(x) = \frac{1}{2}a, \ -a \leq x \leq a; f(x) = 0,$ elsewhere $M(t : X) = (\sinh at) / at$. Also, show that the moments of even order are given by

$$\mu_{2n} = a^{2n}/(2n + 1).$$

3*. Find $M_X(t)$ when the c.d.f. of X is

$$F_X(x) = 0, \ x < 0; \ F_X(x) = 1 - \tfrac{1}{2}e^{-x}, x \geq 0.$$

4*. If the moments of a variate X are defined by $\mathrm{E}(X^r) = 0.6, r = 1, 2, 3, \ldots,$ show that $P(X = 0) = 0.4$, $P(X = 1) = 0.6$, $P(X \geq 2) = 0$.

5*. Let $m(t) = \mathrm{E}(e^{tX}), \mathrm{E}(X) = \mu$ and Var $(X) = \sigma^2$. Let $M(t) = \mathrm{E}(e^{tY}) = e^{c[m(t) - 1]}$ ...(1)
Find the mean and variance of Y in terms of μ and σ^2.

6*. Variate X has the mixed distribution :

$f(x) = \frac{1}{2}\lambda^{-1}e^{-x/\lambda}, 0 < x < \infty;$ omitting the integers; $f(x) = 0$, otherwise.

$$p(x) = \frac{1}{2} e^{-\lambda} \lambda^x / x!, \quad x = 0, 1, 2, \ldots$$

Find m.g.f. Hence or otherwise obtain Var (X).

7*. A r.v. X follows $N(0, 1)$ distribution with probability p and $N(1, 1)$ distribution with probability $q(= 1 - p)$. Find Var (X).

8-20. *Joint Moment Generating Function*

Definition. The joint m.g.f. of random variables : $(X_2, X_2, \ldots, X_n)$ is defined by

$$M_X(t_1, t_2, \ldots, t_n) = \mathsf{E}(\exp[t_1 X_1 + t_2 X_2 + \ldots + t_n X_n)] \qquad \ldots(1)$$

provided the expectaion exists for all values of $t_1, t_2, \ldots, t_n$ such that $-h < t_k < h$ for some $h > 0$, $k = 1, 2, 3, \ldots, n$ and $X = (X_1, X_2, \ldots, X_n)$.

From (1), we obtain the joint raw (i.e. simple) moments by differentiation

$$\mathsf{E}(X_i^r) = \left[\frac{\partial^r M(t_1, t_2, \ldots, t_n)}{\partial X_i^r}\right]_{t=0}, \mathsf{E}(X_i^r X_j^s) = \left[\frac{\partial^{r+s} M(t_1, \ldots, t_n)}{\partial X_i^r \, \partial X_j^s}\right]_{t=0}$$

where $t = (t_1, t_2, \ldots, t_n)$; $t = 0$ means $t_1 = 0, t_2 = 0, \ldots, t_n = 0$.

8-21. *Joint m.g.f. and Mixed moments for Two Variates*

The joint m.g.f. of X and Y is : $M_{X,Y}(t_1, t_2) = \mathsf{E}[e^{t_1 X + t_2 Y}]$.

$$\therefore \quad M_{X,Y}(t_1, 0) = \mathsf{E}(e^{t_1 X}) = M(t_1 : X) \qquad M_{X,Y}(0, t_2) = \mathsf{E}(e^{t_2 Y}) = M(t_2 : Y)$$

Thus, marginals (individual) m.g.f.'s can be obtained from the joint m.g.f.'s by setting $t_2 = 0$ or $t_1 = 0$. We shall write $M_X, Y(t_1, t_2) \equiv M(t_1, t_2)$; $M(t_1; X) \equiv M(t_1)$, $M(t_2 : Y) = M(t_2)$, etc.

Moments. Let $M_c(t_1, t_2) = \mathsf{E}\{e^{t_1(X-a) + t_2(Y-b)}\} = \mathsf{E}\{e^{t_1(X-a)} . e^{t_2(Y-b)}\}$

$$= \mathsf{E}\left\{\sum_{r=0}^{\infty} \frac{t_1^r (X-a)^r}{r!} . \sum_{s=0}^{\infty} \frac{t_2^s (Y-b)^s}{s!}\right\} = \mathsf{E}\left\{\sum_r \sum_s \frac{t_1^r t_2^s}{r!s!} (X-a)^r (Y-b)^s\right\}$$

$$= \sum_{r=0}^{\infty} \sum_{s=0}^{\infty} \frac{t_1^r . t_2^s}{r!s!} \mathsf{E}(X-a)^r (Y-b)^s = \sum_{r,s}^{\infty} \frac{t_1^r . t_2^s}{r!s!} \mu'_{r,s}.$$

Thus, $\mathsf{E}[(X-a)^r (Y-b)^s] = \text{Coeff. of } (t_1^r t_2^s / r!s!) \text{ in } M_c(t_1, t_2)$

In particular : $\mathsf{E}(X^r Y^s) = \text{Coeff. of } (t_1^r t_2^s / r!s!) \text{ in } M(t_1, t_2)$

$$\mathsf{E}[(X - \mu_x)^r (Y - \mu_y)^s] = \text{Coeff. of } (t_1^r t_2^s / r!s!) \text{ in } M_0(t_1 \; t_2).$$

Cor. Assume differentiation under expectation sign (DUES) :

$$\therefore \quad \frac{\partial^{r+s}}{\partial t_1^r \, \partial t_2^s} M_c(t_1, t_2) = \mathsf{E}\left\{\frac{\partial^{r+s}}{\partial t_1^r \, \partial t_2^s} e^{t_1(X-a) + t_2(Y-b)}\right\} = \mathsf{E}\{(X-a)^r (Y-b)^s e^{t_1(X-a) + t_2(Y-b)}\}$$

So :

$$\boxed{\frac{\partial^{r+s}}{\partial^r t_1 \, \partial^s t_2} M_c(0, 0) = \mathsf{E}[(X-a)^r (Y-b)^2].}$$

8-22. Factor Theorem for Independent Variates

Two jointly distributed variates X and Y are independent iff, for some $h > 0$

$$M(t_1, t_2) = M(t_1)\, M(t_2) \;\forall\; t_1, t_2 \text{ with } -h < t_1, t_2 < h.$$

Proof. If X and Y are independent, then

$$M(t_1, t_2) = \mathrm{E}(e^{t_1 X + t_2 Y}) = \mathrm{E}[e^{t_1 X} e^{t_2 Y}] = \mathrm{E}[e^{t_1 X}]\,\mathrm{E}[e^{t_2 Y}] \qquad [\because\; e^{t_1 X} \text{ and } e^{t_2 Y} \text{ are indep.}]$$

i.e. $\quad M(t_1, t_2) = M(t_1, 0)\, M(0, t_2).$ ...(1)

Thus, the stochastic independence of X and Y implies that the joint m.g.f. *factors* into the product of the marginal i.e. individual m.g.f.'s.

Converse. Given result (1), we need to prove that X and Y are independent.

Let f, f_1, f_2 be the joint p.d.f. and marginal p.d.f.'s of X and Y, supposed continuous. Now X, Y have unique m.g.f.'s given by

$$M(t_1, 0) = \int_{-\infty}^{\infty} f_1(x)\, e^{t_1 X}\, dx\,; \quad M(0, t_2) = \int_{-\infty}^{\infty} f_2(y)\, e^{t_2 Y}\, dy.$$

$$\therefore \quad M(t_1, 0)\, M(0, t_2) = \int_{-\infty}^{\infty} f_1(x)\, e^{t_1 x} dx \int_{-\infty}^{\infty} f_2(y)\, e^{t_2 y} dy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_1(x)\, f_2(y)\, e^{(t_1 X + t_2 Y)}\, dx\, dy \quad ...(2)$$

Using relation (1), Eq. (2) can be rewritten as

$$M(t_1, t_2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_1(x)\, f_2(y) \exp(t_1 x + t_2 y)\, dx\, dy. \quad ...(3)$$

But $M(t_1, t_2)$ is the joint m.g.f. of X and Y, so

$$M(t_1, t_2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x, y) \exp(t_1 x + t_2 y)\, dx\, dy \quad ...(4)$$

The uniqueness of m.g.f. implies that (3) and (4) must yield

$$f(x, y) = f_1(x)\, f_2(y) \;\Rightarrow X \text{ and } Y \text{ are } \textit{independent.}$$

The proof, when X and Y are discrete, follows by using summation instead of integration.

Remark. The result can be extended to any number of variates.

Example : Let X, U, V be indep. variates such that U has the Bernoulli distribution

$$P\{U = 0\} = p,\; P\{U = 1\} = q,\; p + q = 1,\; 0 < p < 1,\; V \sim \text{Expo}\,(\lambda).$$

Let $Y = pX + UV$ and suppose that X and Y are identically distributed. Show that,

(a) The common distribution of X and Y is Expo (λ), (b) Corr $(X, Y) = p$,

(c) Regression of Y on X is linear.

Solution. Let $Z = UV$; Now using Double-E Rule, we get

$$M(t : UV) = \mathrm{E}(e^{tUV}) = \mathrm{E}\{\mathrm{E}(e^{(tUV)} \mid u)\} = \mathrm{E}[1 - (tU/\lambda)]^{-1} \qquad [\text{m.g.f. of Expo}(\lambda)]$$

$$M(t : Z) = p\,.(1 - 0)^{-1} + q\,.(1 - t/\lambda)^{-1} = p + q\lambda/(\lambda - t)^{-1} = (\lambda - pt)/(\lambda - t). \quad ...(1)$$

(a) $M(t : Y) = M(t : pX + Z) = M(pt : X)\,.\,M(t : Z)$

$$\therefore \quad \frac{M(t : Y)}{M(pt : X)} = M(t : Z) = \frac{\lambda - pt}{\lambda - t} = \frac{1 - pt/\lambda}{1 - t/\lambda} = \frac{(1 - t/\lambda)^{-1}}{(1 - pt/\lambda)^{-1}}. \qquad [\text{by (1)}] \quad ...(2)$$

We observe that $M(pt : X) = M(t : pX)$ and if $X \sim$ expo (λ) then $pX \sim$ expo (λ/p). Thus, if λ is changed to λ/p, the Denominator equals Numerator in (2). It follows that since X and Y are identically distributed; $X \sim$ expo (λ).

(b) $\text{Cov}(X, Y) = \text{Cov}(X, pX + Z) = p\,\text{Cov}(X, X) + \text{Cov}(X, Z) = p\sigma_X^2 + 0.$ [Ind (X, Z)]

$\therefore$ Corr $(X, Y) = \sigma_{XY} / \sigma_X\sigma_Y = p\sigma_X^2 / \sigma_X^2 = p.$ [$\sigma_X = \sigma_Y$ as X & Y are i.d.]

(c) $E(Y \mid X = x) = E\{(pX + Z) \mid X = x\} = pE(X \mid x) + E(Z \mid x) = px + E(Z).$ [Ind (X, Z)]

Since $M_Z'(t) = (q/\lambda)(1 - t/\lambda)^{-2}, E(Z) = M_Z'(0) = q/\lambda.$

$\therefore$ $E(Y \mid X = x) = px + (q/\lambda).$ [Regression of Y on X is linear]

8-23. The *n*-variate Normal Distribution

Let $(X_1, X_2, \ldots, X_n)$ be an n-variate vector defined on the sample space S. Let **X** be an n-dimensional random vector and **X** be an n-dim vector, both expressed as $n \times 1$ matrices, i.e. column vectors :

$$\mathbf{X} = [X_1, X_2, \ldots, X_n]^T, \ \mathbf{x} = [x_1, x_2, \ldots, x_n]^T \qquad [T = \text{transpose}]$$

The n-variate vector $(X_1, X_2, \ldots, X_n)$ is called an n-variate *Normal* random vector, if its joint p.d.f. is given by

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{(\sqrt{2\pi})^n (\det K)^{1/2}} \exp\left[-\tfrac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T (K^{-1}) (\mathbf{x} - \boldsymbol{\mu})\right]$$

where $\boldsymbol{\mu} = (\mu_1, \mu_2, \ldots, \mu_n)$ is the *vector mean* $[\mu_j = E(X_j)]$ and K is the *covariance matrix* $[\sigma_{ij}]$, $\sigma_{ij} = \text{Cov}(X_i, X_j)$ and det K = determinant of the matrix K. We understand the notation :

$$f_{\mathbf{X}}(\mathbf{x}) = f_{X_1, \ldots, X_n}(x_1, \ldots, x_n)$$

Alternate Approach :

Let $Z_j \sim N(0, 1), 1 \le j \le n$ be indep. variates. If, for some constant μ_i and a_{ij}, $1 \le i \le m$,

$$X_i = \mu_i + a_{i1}Z_1 + a_{i2}Z_2 + \ldots + a_{in}Z_n, \qquad i = 1, 2, 3, \ldots, m.$$

then the r.v.s. $X_1, X_2, \ldots, X_m$ are said to possess *multivariate normal distribution.* Since the linear combination of indep. normal variates is a normal variate, it follows that each X_i is a normal variate. Further,

$$E(X_i) = \mu_i, \text{ var }(X_i) = \text{Var}\left(\mu_i + \sum_{i=1}^{n} a_{ij}Z_i\right) = \sum_{j=1}^{n} a_{ij}^2. \qquad [\text{Var}(Z_j) = 1, E(Z_j) = 0]$$

Since Cov. $(Z_r, Z_s) = E(Z_rZ_s) = 1$, if $r = s$; $E(Z_rZ_s) = 0$, if $r \ne s$. So Cov $(Z_r, Z_s) = \delta_{rs}$.

$$\therefore \sigma_{ij} = \text{Cov}(X_i, X_j) = \text{Cov}\left((\mu_i + \sum_{r=1}^{n} a_{ir}Z_r, \ \ \mu_j + \sum_{s=1}^{n} a_{is}Z_s\right) = \text{Cov}\left(\sum_{r=1}^{n} a_{ir}Z_r, \ \sum_{s=1}^{n} a_{is}Z_s\right)$$

$$= \sum_{x=0}^{n}\sum_{x=0}^{n} a_{ir}a_{is} \text{ Cov}(Z_r, Z_s) = \sum_r \sum_s a_{ir}a_{js}\,\delta_{rs} = \sum_{r=1}^{n} a_{ir}\,a_{jr}. \qquad \ldots(1)$$

To determine the joint distribution of $X_1, \ldots, X_m$ we find their joint m.g.f.. Now

$$Y = t_1X_1 + t_2X_2 + \ldots + t_mX_m = t_1(\mu_1 + \Sigma a_{1j}Z_j) + t_2(\mu_2 + \Sigma a_{2j}Z_j) + \ldots + t_m(\mu_m + \Sigma a_{mj}Z_j)$$

$$= \Sigma\mu_it_i + (\Sigma t_ia_{i1})Z_1 + (\Sigma t_ia_{i2})Z_2 + \ldots + (\Sigma t_ia_{in})Z_n, \quad 1 \le i \le m.$$

We use prime $'$ to indicate transposition.

Obviously, $E(Y) = E(\Sigma t_i X_i) = \Sigma t_i \mu_i = t\mu'$ (row vector × column vector)

$$\text{Var}(Y) = \text{Var}(\Sigma t_i X_i) = \text{Var}[\Sigma \mu_i t_i + (\Sigma t_i a_{i1})Z_1 + \ldots + (\Sigma t_i a_{in}) Z_n]$$

$$= \sum_{s=1}^{n}\left(\sum_{i=1}^{m} t_i\, a_{is}\right)^2 = \sum_{s=1}^{n}\left(\sum_{i=1}^{m} t_i a_{is}\right)\left(\sum_{k=1}^{m} t_k\, a_{ks}\right) = \sum_{i=1}^{m}\sum_{k=1}^{m} t_i\, t_k \left(\sum_{s=1}^{n} a_{is}\, a_{ks}\right)$$

$$= \sum_{i=1}^{m}\sum_{k=1}^{m} t_i\, t_k\, \sigma_{ik} = t(\sigma_{ij})\, t'. \qquad \text{[by (1)]} \qquad [(\sigma_{ij}) \text{ is } m \times m \text{ matrix}]$$

Recall : $Y \sim N(\mu, \sigma^2)$, then $E(e^Y) = e^{\mu + (\sigma^2/2)}$, [$t = 1$ in m.g.f. of $N(\mu, \sigma^2)$]

$$\therefore \qquad M(t_1, \ldots, t_m) = E[\exp(t_1 X_1 + \ldots + t_m X_m)] = E(e^Y) = \exp\left[\Sigma t_i\, \mu_i + \frac{1}{2}\sum_{i=1}^{m}\sum_{k=1}^{m} t_i t_k \sigma_{ik}\right]$$

$$= \exp\left\{t\mu' + \frac{t(\sigma_{ij})t'}{2}\right\}$$

where $t = (t_1, \ldots, t_n)$, prime indicates transpose and (σ_{ij}) is *dispersion* (covariance) matrix.

Note. $f(x) = \{(2\pi)^m\, |\sigma_{ij}|\}^{-1/2} \exp[-\frac{1}{2}(x-\mu)[(\sigma_{ij})]^{-1}(x-\mu)'], -\infty < x_i < \infty,\ 1 \le i \le m.$

This shows that the joint distribution of $(X_1, \ldots, X_m)$ is completely determined from the knowledge of μ_i and σ_{ik}; $i, k = 1, 2, \ldots, m$.

Cor. **For bivariate normal**

$$M(t_1, t_2) = \exp[(\mu_1 t_1 + \mu_2 t_2) + \tfrac{1}{2}(\sigma_1^2 t_1^2 + 2\rho\sigma_1\sigma_2\, t_1 t_2 + \sigma_2^2 t_2^2)]. \qquad [\sigma_{12} = \sigma_1\sigma_2\rho]$$

Chapter 20 is devoted to the study of bivariate normal distribution.

Example : Let $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ be n-variate Normal vector. Show that if σ_{ij} = Cov $(X_i, X_j) = 0$, $\forall i, j$ $(i \ne j)$, then $X_1, X_2, \ldots, X_n$ are independent.

Solution. Given $\sigma_{ij} = 0$ if $i \ne j, \sigma_{ii} = \sigma_i^2$ $(i = j)$ the covariance matrix K and its inverse are immediate

$$K = \begin{bmatrix} \sigma_1^2 & 0 \ldots 0 \\ 0 & \sigma_2^2 \ldots 0 \\ \ldots & \ldots\ldots \\ 0 & 0 \ldots \sigma_n^2 \end{bmatrix}, \quad K^{-1} = \begin{bmatrix} (\sigma_1^2)^{-1} & 0 \ldots 0 \\ 0 & (\sigma_2^2)^{-1} \ldots 0 \\ \ldots & \ldots\ldots \\ 0 & 0 \ldots (\sigma_n^2)^{-1} \end{bmatrix}, \quad \det K = \sigma_1^2, \sigma_2^2, \ldots \sigma_n^2$$

$(\mathbf{x} - \boldsymbol{\mu}) = (x_1 - \mu_1, x_2 - \mu_2, \ldots, x_n - \mu_n)$ and thus

$$(\mathbf{x} - \boldsymbol{\mu})^T (K^{-1})(\mathbf{x} - \boldsymbol{\mu}) = \Sigma\left(\frac{x_i - \mu_i}{\sigma_i}\right)^2 \qquad [(\mathbf{x} - \boldsymbol{\mu})^T = \text{row vector}]$$

$$\therefore \ f(x_1, \ldots, x_n) = \frac{1}{(\sqrt{2\pi})^n} \cdot \frac{1}{\sigma_1, \sigma_2, \ldots, \sigma_n} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}\left(\frac{x_i - \mu_i}{\sigma_i}\right)^2\right\}$$

Thus, $\qquad f(x_1, x_2, \ldots, x_n) = f_1(x_1)\, f_2(x_2) \ldots f_n(x_n)$ ...(1)

where $\qquad f(x_i) = (\sqrt{2\pi}\, \sigma_i)^{-1}\, e^{-(x_i - \mu_i)^2/\sigma_i^2},\ i = 1, 2, \ldots, n.$

Equation (1) implies that $X_1, X_2, \ldots, X_n$ are independent normal variates.

8-30. Descending Factorial M.G.F.

The factorial m.g.f. of a random variable X is denoted by $W(t : X)$ and is defined by the relation

$$W(t : X) = E[(1 + t)^X] \quad ...(1)$$

Since $(1+t)^x = \Sigma x^{(r)} t^r / r!$ where $x^{(r)} = x(x-1)(x-2)\ldots(x-r+1)$ we can rewrite (1) as

$$W(t : X) = E(\Sigma X^{(r)} t^r / r! = E[X^{(r)}] t^r / r! = \Sigma \mu'_{(r)} t^r / r! \quad ...(2)$$

where $\mu'_{(r)} = \Sigma[X^{(r)}]$ is the rth-order simple factorial moment. Eq. (2) explains why result (1) is given the name of factorial m.g.f.

We combine (1) and (2)

$$\boxed{W(t) = E[(1+t)^X] = \Sigma \mu'_{(r)} t^r / r!} \quad ...(3)$$

Alternate. Sometimes, we define descending factorial m.g.f. via $G_X(t) = E(t^X)$.

Here, $G'(t) = E(Xt^{X-1}), G''_X(t) = E(X^{(2)} t^{X-2}), \ldots, G^k_X(t) = E(X^{(k)} t^{X-k})$

Putting $t = 1$, we get

$$G^k_X(1) = \mu'_{(k)} = E(X^{(k)}), k = 1, 2, 3, \ldots$$

Remark. Sometimes, the simple moments $E(X^r)$ can be most beautifully determined from the factorial simple moments $E[X^{(r)}]$. As such, the function $W(t : X)$ is of sufficient interest.

8-31. *Factorial moments via factorial m.g.f. for some Discrete Distributions*

1. *Binomial distribution.* $X \sim$ bin (n, p)

$$W(t : X) = E(1+t)^X = \sum_{x=0}^{n} \binom{n}{x} q^{n-x} [p(1+t)]^x = [q + p(1+t)]^n = (1 + pt)^n \quad ...(1)$$

Using (1) $W(t : X) = \sum_{r=0}^{\infty} \mu'_{(r)} \frac{t^r}{r!} = \sum_{r=0}^{\infty} \frac{n^{(r)}}{r!} (pt)^r \Rightarrow \mu'_{(r)} = n^{(r)} p^r$. ...(2)

Thus $\mu'_1 = E(X) = np$, $\mu'_{(2)} = EX(X-1) = n(n-1)p^2 \Rightarrow E(X^2) = n^2p^2 + npq$. Var $(X) = npq$. ...(3)

2. *Poisson distribution.* $X \sim$ Pois (λ)

$$W(t : X) = E(1+t)^X = \sum_{r=0}^{\infty} \frac{e^{-\lambda}\lambda^x}{x!} (1+t)^x = e^{-\lambda} \sum_{r=0}^{\infty} \frac{[\lambda(1+t)]^x}{x!} = e^{-\lambda} e^{\lambda(t+1)} = e^{\lambda t}. \quad ...(1)$$

Using (1), $W(t : X) \equiv \sum_{r=0}^{\infty} \mu'_{(r)} \frac{t^r}{r!} = \sum_{r=0}^{\infty} \frac{(\lambda t)^r}{r!} \Rightarrow \mu'_{(r)} = \lambda^r$. ...(2)

Thus, $\mu'_1 = E(X) = \lambda$, $\mu'_{(2)} = E[X(X-1)] = \lambda^2 \Rightarrow E(X^2) = \lambda + \lambda^2$; Var $(X) = \lambda$ ...(3)

3. *N-B distribution.* $X \sim NB(k\ ; p)$.

$$W(t : X) = E(1+t)^X = \sum_{x=0}^{\infty} \binom{k+x-1}{x} p^k q^x (1+t)^x = p^k [1 - q(1+t)]^{-k}$$

$$= p^k (p - qt)^{-k} = \left(1 - \frac{qt}{p}\right)^{-k} \quad ...(1)$$

Using (1), $W(t:X)=\sum_{r=0}^{\infty}\mu'_{(r)}\dfrac{t^r}{r!}=\sum_{r=0}^{\infty}\binom{k+r-1}{r}\left(\dfrac{q}{p}\right)^r t^r \Rightarrow \mu'_{(r)}=(k+r-1)^{(r)}\left(\dfrac{q}{p}\right)^r$ ...(2)

So $\mu'_1=kq/p, \mu'_{(2)}=(k+1)k\,(q/p)^2 \Rightarrow \mathsf{E}(X^2)=k^2\dfrac{q^2}{p^2}+\dfrac{kq}{p^2}$; Var $(X)=\dfrac{kq}{p^2}$ ...(3)

8-32. Linear Combination Property

If X and Y are two *independent* variates and a, b, c any constant, then

$$W(t:aX+bY+c)=(1+t)^c\,W(t:aX)\,W(t:bY).$$

Proof. $W(t:aX+bY+c)=\mathsf{E}\{(1+t)^{aX+bY+c}\}=\mathsf{E}\{(1+t)^c\,.\,(1+t)^{aX}\,.\,(1+t)^{bY}\}$ [by Def]

$=(1+t)^c\,.(1+t)^{aX}\,.(1+t)^{bY}=(1+t)^c\,W(t:aX)\,.\,W(t:bY).$ [by Lin E, indep. & Def.]

8-33. Relation between Simple and Central Factorial Moments

The factorial m.g.f. about $X=0$ and $X=\mu$ are connected by

$$W(t:X-\mu)=(1+t)^{-\mu}\,W(t:X)$$

$$\therefore\ \sum_r \mu_{(r)}\frac{t^r}{r!}\left(\sum_j\frac{(-\mu)^{(j)}t^j}{j!}\right)\left(\sum_j \mu'_{(s)}\frac{t^s}{s!}\right)=\sum_s\sum_j\left(\frac{t^r}{r!}\right)\frac{r!}{s!\,j!}(-\mu)^{(j)}\,\mu'_{(s)}\quad [s=r-j]$$

Comparing the coefficient of $t^r/r\,!$ on both sides, we get

$$\mu_{(r)}=\sum_{j=0}^{r}\frac{r!}{(r-j)!\,j!}(-\mu)^j\,\mu_{(r-j)}=\sum_{j=0}^{r}\binom{r}{j}\mu'_{r-j}\,(-\mu)^j.$$

In particular taking $r=2, 3, 4$ we obtain

$$\mu_{(2)}=\mu'_{(2)}-\mu^2+\mu_0 \qquad [\mu'(1)=\mu]$$

$$\mu_{(3)}=\mu'_{(3)}-3\mu'_{(2)}\,\mu+2\mu^3-2\mu.$$

$$\mu_{(4)}=\mu'_{(4)}-4\mu'_{(3)}\,\mu+6\mu'_{(2)}\,\mu(\mu+1)-3\mu(\mu^3+2\mu^2-\mu-2).$$

Exercise 8(a)

1. Can $M(t)=\exp(t^2/2)$ be the m.g.f. of some random variable ? [No]
2. If $\mu'_r=3^r$, $r=0, 1, 2, \ldots$, obtain the p.m.f. of X. [$P(X=3)=1$]
3. Obtain the m.g.f. of the random variable for which

   (i) $\mu'_r=r\,!.\ r=0, 1, 2, \ldots$ (ii) $\mu'_r=(r+1)\,!\,2^r,\ r=0, 1, 2, \ldots$ [**Ans.** $(1-t)^{-1}, (1-2t)^{-2}$]
4. Show that if $\overline{X}_n$ is the mean of n i.i.d. variates each distributed as X, then $M(t:\overline{X}_n)=[M(t/n:X)]^n$.
5. The variate X has the Laplace law : $f(x)=\frac{1}{2}\theta^{-1}\exp(|x-\theta|/\theta), -\infty<x<\infty.$

   Find the m.g.f. of X and hence show that $\mathsf{E}(X)=\theta$ and Var $(X)=2\theta^2$. Also obtain $\mathsf{E}(X^n e^{x/2})$.
6. A variate X has p.d.f. : $f(x)=(\frac{1}{2})^x, x=1, 2, 3, \ldots$ Find its m.g.f. and variance. Show also $\mathsf{E}(X^n)=\Sigma(k^n/2^k),\ 1\le k<\infty.$

7. A variate X has p.d.f. : $f(x) = ab^x, x = 0, 1, 2, 3, ..., a + b = 1, a > 0, b > 0$. Find the m.g.f. of X and deduce that $m_2 = m_1(1 + 2m_1)$, where m_1, m_2 are the first two simple moments.

8. Obtain the m.g.f. of the variate X having p.d.f.
$f(x) = x, 0 \le x < 1; f(x) = 2 - x, 1 \le x < 2; f(x) = 0$, elsewhere.
Show that $\mu_2 = 1/6, \mu_3 = 0$ and $\mu_4 = 1/15$.

9. A variate X has the p.d.f. given by
(a) $f(x) = 1/3, -1 < x < 2; f(x) = 0$, elsewhere.
(b) $f(x) = 1/2\sqrt{x}, 0 < x < 1; f(x) = 0$ elsewhere.
Find the m.g.f. and hence the mean and variance.

$$\left[\textbf{Ans. } (b)\ M(t) = \sum_{k=0}^{\infty} \frac{1}{(2k+1)} \frac{t^k}{k!}, \mathrm{E}(X) = \frac{1}{3} \text{ Var }(X) = \frac{4}{45}\right]$$

10. Let X be a mixed variate such that
$f(x) = c, 0 < x < 1, 0 < c < 1; f(x) = 0$ elsewhere.
$P(x) = \binom{x+r-1}{r-1} p^r (1-p)^x (1-c), x = 0, 1, 2, ...$ (r an integer)
Show that $M(t) = c(e^t - 1)t^{-1} + (1-c)[p/(1-qe^t)]^r$, and find the mean and the variance.

11. Let X be a mixed variate such that for $a > 0, \lambda > 0, 0 < k < 1$;
$f(x) = kx^{a-1}e^{-\lambda x}\lambda^a / \Gamma(a), 0 < x < \infty$, omitting the intergers 1, 2, 3, ..., $f(x) = 0$, otherwise.
$P(x) = (1-k)\binom{n}{x} p^x (1-p)^{n-x}, x = 0, 1, 2, ..., n, 0 < p < 1$.
Show that $M(t) = k(1 - t/\lambda)^{-a} + (1-k)(q + pe^t)^n$ and find the mean and variance.

12. For the p.m.f. : $P(X = x) = a_x\theta^x/f(\theta), 0 \le x < \infty, \theta > 0$ where $a_x \ge 0$ and $f(\theta) = \Sigma_x a_x \theta^x$. Show that $M(t) = f(\theta e^t)/f(\theta)$.

13. The m.g.f.'s of variates X and Y are expressed as power series :
$$M(t : X) = \sum_{r=0}^{\infty} t^r, |t| < 1; \quad M(t : Y) = \sum_{r=0}^{\infty} \sum_{k=0}^{r} \frac{2^{r-k}}{k!} t^r, |t| < \tfrac{1}{2}.$$
(a) Find $\mathrm{E}(X^n)$ and $\mathrm{E}(Y^n)$, $n = 0, 1, 2, ...$ (b) How are the variates X and Y related ?

14. Independent variates X and Y have the p.m.f. :
$P(=x) = x/6, x = 1, 2, 3, ; P(Y = y) = (y+2)/10, y = -1, 2, 3$.
Using m.g.f.'s, find the distributions of $S = X + Y$ and $D = X - Y$.
[**Ans.** $f(0) = p, f(1) = 2p, f(2) = 3p, f(3) = 4p, f(4) = 13p, f(5) = 22p, f(6) = 15p, p = 1/60$.
$g(-2) = 5p, g(-1) = 14p, g(0) = 23p, g(1) = 12p, g(2) = p, g(3) = 2p, g(4) = 3p$]

15. If the joint m.g.f. of (X, Y) is (i) $M(t_1, t_2) = \exp [\frac{1}{2}(t_1^2 + t_2^2)]$, (ii) $M(t_1, t_2) = \exp [(t_1 + t_2)/2]$, what is the distribution of Y ?

16. Continuous variates X and Y are jointly distributed as
$f(x, y)\, e^{-y}, 0 < x < y < \infty; f(x, y) = 0$, elsewhere.
Show that the $M(t_1, t_2) = [(1 - t_2)(1 - t_1 - t_2)]^{-1}$ and hence prove that X and Y are not independent.

17. Let X, Y, Z have the joint p.m.f.

$f(x, y, z) = \frac{1}{4}$, $(x, y, z) \in [(1, 0, 0), (0, 1, 0), (0, 0, 1), (1, 1, 1)]$, $f(x, y, z) = 0$, elsewhere. Using m.g.f. show that X, Y, Z are pairwise indep. but not mutually independent.

18. Define the m.g.f. of $Y \mid X = x$. Does $M(t : Y) = \mathrm{E}[M(t : Y \mid X = x)]$?

19. A variate X is known to have the distribution : $dF = k[1 + (x/a)]^{m-1} e^{-mx/a} dx$, $-a \le x < \infty$. Find the constant k and determine the linear relation between β_1 and β_2 of this distribution. Also prove that $M(t) = e^{-at}[1 - (at/m)]^{-m}$.

8-40. *Cumulant Generating Function*

The cumulant generating function (notation : c.g.f.) of a random variable X, denoted by $K(t : X)$ or $K_X(t)$, is defined by the relation.

$$K(t : X) = \log M(t : X) = \log \Sigma \mu_r' (t^r / r!), \ 0 \le x < \infty \qquad \text{...(1)}$$

provided the resulting series on the right side in (1) is convergent. If the Taylor's series for $K(t)$ be

$$K(t) = k_1 t + k_2 t^2 / 2! + k_3 t^3 / 3! + \ldots \qquad \text{...(2)}$$

then the coefficient k_r is called the rth cumulant of random variable X; $r = 1, 2, \ldots$

If we differentiate (2) w.r.t. 't' we readily obtain

$$[D^r K(t)]_{t=0} = K^{(r)}(0) = k_r . \qquad (D = d/dt) \qquad \text{...(3)}$$

8-41. *Effect of Linear Transformations on K(t : X)*

Since $\quad M(t : a + bX) = e^{at} M(bt : X)$, taking its logarith yields

$$= \ln M(t : a + bX) = at + \ln M(bt : X)$$

$$\Sigma k_r'(t^r / r!) = at + \Sigma k_r (bt)^r / r!, \ (r = 1, 2, 3, \ldots)$$

Comparing coefficients of t^r on both sides we obtain

$$k_1' = a + bk_1; \ k_r' = b^r k_r, (r \ge 2). \qquad \text{...(1)}$$

Thus, all cumulants except k_1 are unaltered by the change of origin (factor a) and rth cumulant of $Y = (a + bX)$ is b^r times the rth cumulant of X. Thus, a linear map possesses *semi-invariant* property. Eq. (1) governs the effect of change of origin and scale on $K(t)$.

8-42. *Cumulants in terms of Central Moments*

Since $M(t : X - \mu) = e^{-\mu t} M(t : X)$, taking its logarithm yields

$$\ln M(t : X) = \ln M(t : X - \mu) + \mu t.$$

Inserting power series for m.g.f. and $K(t)$, we get

$$\sum_{r=1}^{\infty} k_r \frac{t^r}{r!} = \mu t + \ln \left\{ 1 + \sum_{r=1}^{\infty} \mu_r \frac{t_r}{r!} \right\} \qquad [\mu_1 = 0]$$

$$= \mu t + \left\{ \sum_{2}^{\infty} \mu_r \frac{t^r}{r!} \right\} - \frac{1}{2} \left\{ \sum_{2}^{\infty} \mu_r \frac{t^r}{r!} \right\}^2 - \frac{1}{3} \left\{ \sum_{2}^{\infty} \mu_r \frac{t^r}{r!} \right\}^3 - \frac{1}{4} \left\{ \sum_{2}^{\infty} \mu_r \frac{t^r}{r!} \right\}^4 + \ldots$$

The first few relations, using k_r = Coeff. of $(t^r/r\,!)$, are obtainable :

$$\boxed{k_1 = \mu, k_2 = \mu_2, k_3 = \mu_3,\ k_4 = \mu_4 - 3\mu_2^2,\ k_5 = \mu_5 - 10\mu_3\mu_2,\ \ k_6 = \mu_6 - 15\mu_4\mu_2 - 10\mu_3^2 + 30\mu_2^3}$$

8-43. *Simple Moments in Terms of Cumulants*

We invert the definition : $K(t) = \log M(t : X)$ to obtain $M(t : X) = e^K(t)$. Inserting power series expansion of $M(t)$ and $K(t)$, this gives

$$\sum_{r=0}^{\infty} \mu_r' \frac{t^r}{r!} = \exp\left[\sum_{r=0}^{\infty} k_r \frac{t^r}{r!}\right] = 1 + \left(\Sigma k_r \frac{t^r}{r!}\right) + \frac{1}{2}\left(\Sigma k_r \frac{t^r}{r!}\right)^2 + \frac{1}{6}\left(\Sigma k_r \frac{t^r}{r!}\right)^3 + \ldots \qquad \ldots(1)$$

The first few relations, obtained by comparing coeff. of $t^r/r\,!$, are

$$\mu_1' = k_1,\ \mu_2' = k_2 + k_1^2,\ \mu_3' = k_3 + 3k_1k_2 + k_1^3;\ \mu_4' = k_4 + (3k_2^2 + 4k_1k_3) + 6k_1^2k_2 + k_1^4.$$

8-44. General Linear Property

$$k_r(aX + bY) = a^r k_r(X) + b^r k_r(Y) \qquad \ldots(1)$$

where X and Y are *independent; a, b* are constants.

Proof. $M(t : aX + bY) = M(at : X)\, M(bt : Y)$ $\qquad$ ($\because$ X, Y are indep.)

$\therefore$ $\ln M(t : aX + bY) = \ln M(at : X) + \ln M(bt : Y)$

or $K(t : aX + bY) = K(at : X) + K(bt : Y)$. $\qquad \ldots$(i)

We now equate the Coeff. of $t^r/r\,!$ on both sides of (i) and Eq. (1) emerges readily.

Extension. If X_1, X_2, ..., X_n are independent variates, then

$$k_r(t : a_1X_1 + \ldots + a_nX_n) = a_1^r\, K_r(X_1) + \ldots + a_n^r K_r(X_n). \qquad \ldots(2)$$

The proof follows by mathematical induction.

Cor. If $\overline{X} = (X_1 + X_2 + \ldots + X_n)/n$, then putting $a_i = 1/n$ in (2), we get

$$k_r(\overline{X}) = n^{-r}[k_r(X_1) + k_r(X_2) + \ldots + k_r(X_n)] \qquad \ldots(3)$$

$$= (1/n)^{r-1} k_r(X_1) \qquad \text{[if } X_i \text{ are i.i.d.]} \qquad \ldots(4)$$

8-45. *Excessive-Kurtosis Theorem*

The excess of kurtosis of Sampling distributions of mean $\overline{X}$ is $(1/n)$ times the excess of kurtosis of the given population.

Proof. If γ_2 denotes the excess of kurtosis then by 8-44(4).

$$\gamma_2(\overline{X}) = \frac{k_4(\overline{X})}{[k_2(\overline{X})]^2} = \frac{k_4/n^3}{(k_2/n)^2} = \frac{1}{n}\frac{k_4(X)}{[k_2(X)]^2} = \frac{1}{n}\gamma_2(X).$$

$\therefore$ Excess of kurtosis of $\overline{X} = (1/n)$ [Excess of kurtosis of population X]

8-46. *Worked-out Problems*

***Example* 1.** If $\mu_r' = \mathrm{E}(X^r)$ and k_j is the jth cumulant show that

$$\mu_r' = \sum_{j=1}^{\infty} \binom{r-1}{j-1} k_j \mu_{r-j}'.$$

Solution. Insert power series expansions of $M(t)$ and $K(t)$ in the definition $\ln M(t:X) = K(t:X)$.

$$\therefore \quad \ln\left\{\sum_{n=0}^{\infty} \frac{t^n}{n!}\mu'_n\right\} = \sum_{j=1}^{\infty} k_j \frac{t^j}{j!} \quad \ldots(1)$$

We assume that the power series invovled are uniformly convergent, so that term by term differentiation is permissible. Differentiating (1),

$$\left\{\sum_{n=1}^{\infty} \frac{t^{n-1}}{(n-1)!}\mu'_n\right\}\Big/\left\{\sum_{n=0}^{\infty} \frac{t^n}{n!}\mu'_n\right\} = \sum_{j=1}^{\infty} k_j \frac{t^{j-1}}{(j-1)!}$$

We now clear the fraction, replace the dummy n by r (only on the L.*H*.S.) and obtain

$$\sum_{r=1}^{\infty} \frac{t^{r-1}\mu'_r}{(r-1)!} = \left[\sum_{j=1}^{\infty} \frac{k_j t^{j-1}}{(j-1)!}\right]\left[\sum_{n=1}^{\infty} \frac{t^{n-1}\mu'_{n-1}}{(n-1)!}\right]$$

$$= \sum_{j=1}^{\infty}\sum_{n=1}^{\infty} \frac{k_j\,\mu'_{n-1}t^{n+j-2}}{(j-1)!\,(n-1)!} \equiv \sum_{n=1}^{\infty}\sum_{j=1}^{\infty} \frac{k_j\mu'_{n-1}\,(r-1)!}{(j-1)!\,(n-1)!}\left(\frac{t^{r-1}}{(r-1)!}\right)(n = r - j + 1)$$

We now equate the coefficients of $t^{r-1}/(r-1)!$ on both sides [use $n = r - j + 1$] to get

$$\mu'_r = \sum_{j=1}^{r} \frac{(r-1)!\,k_j\,\mu'_{n-1}}{(j-1)!\,(r-j)!} = \sum_{j=1}^{r} \binom{r-1}{j-1} k_j\,\mu'_{r-j}$$

Example 2. Prove or disprove : $M(t:X+Y) = M(t:X)\,M(t:Y) \Leftrightarrow (X, Y)$ are independent.

Solution. Consider the joint distribution of X and Y as shown in adjoining table [$p = 1/9$].

$x \downarrow$ / $y \rightarrow$	0	1	2	$P(x)$
0	p	0	$2p$	$3p$
1	$2p$	p	0	$3p$
2	0	$2p$	p	$3p$
$P(y)$	$3p$	$3p$	$3p$	$9p = 1$

Since $\quad f(2, 1) = 2p \neq f_1(2)\,f_2(1) = 9p^2$,

we conclude that X and Y are not independent. Also

$$M_X(t) = E(e^{tX}) = 3p(1 + e^t + e^{2t}) = M_Y(t).$$

Let $Z = X + Y$, and write $p_r = P(Z = r)$. Then

$$p_0 = p,\ p_1 = 2p,\ p_2 = 3p,\ p_3 = 2p,\ p_4 = p.$$

Observe the formation, *e.g.* $p_2 = P(Z = 2) =$

$P\{(0, 2) \uplus (1, 1) \uplus (2, 0)\} = 2p + p + 0 = 3p.$

$M_z(t) = p(1 + 2e^t + 3e^{2t} + 2e^{3t} + e^{4t}) = p(1 + e^t + e^{2t})^2$

Thus $M(t : X + Y) = M(t : X)\,M(t : Y)$, although X and Y are not independent.

Example 3. Let $K(t_1, t_2) = \ln M(t_1, t_2)$. Show that

$$\frac{\partial K(0,0)}{\partial t_1} = E(X),\ \frac{\partial K(0,0)}{\partial t_2} = E(Y),\ \frac{\partial^2 K(0,0)}{\partial t_1^2} = \mathrm{Var}(X),\ \frac{\partial^2 K(0,0)}{\partial t_2^2} = \mathrm{Var}\,(Y);\ \frac{\partial^2 K(0,0)}{\partial t_1 \partial t_2} = \mathrm{Cov}\,(X, Y).$$

Solution. We apply logarithm to central m.g.f. : $m(t_1, t_2) = E[e^{t_1(X-\mu_1)+t_2(Y-\mu_2)}] = e^{-(\mu_1 t_1 + \mu_2 t_2)} M(t_1, t_2)$.

$M(t_1, t_2) = \exp\,(\mu_1 t_1 + \mu_2 t_2)\,m(t_1, t_2) \Rightarrow K(t_1, t_2) = \mu_1 t_1 + \mu_2 t_2 + \ln m(t_1, t_2)$

We write : $m_1 = \dfrac{\partial m}{\partial t_1},\ m_2 = \dfrac{\partial m}{\partial t_2},\ m_{12} = \dfrac{\partial^2 m}{\partial t_1 \partial t_2};\ m_{11} = \dfrac{\partial^2 m}{\partial t_1^2},\ m_{22} = \dfrac{\partial^2 m}{\partial t_2^2}$ and note :

$m_1(0,0) = m_2(0,0) = 0,\ m_{11}(0,0) = \sigma_X^2,\ m_{22}(0,0) = \sigma_Y^2,\ m_{12}(0,0) = \sigma_{12}$.

Now $\dfrac{\partial K}{\partial t_1} = \mu_1 + \dfrac{m_1}{m},\ \dfrac{\partial K}{\partial t_2} = \mu_2 + \dfrac{m_2}{m},\ \dfrac{\partial^2 K}{\partial t_1^2} = \dfrac{mm_{11} - m_1^2}{m^2},\ \dfrac{\partial^2 K}{\partial t_1 \partial t_2} + \dfrac{mm_{12} - m_1 m_2}{m^2}$

$\because\ \partial_1 K(0,0) = \mu_1,\ \partial_2 K(0,0) = \mu_2,\ \partial_{12}K(0,0) = \sigma_{XY},\ \partial_{11}K(0,0) = \sigma_X^2,\ \partial_{22}K(0,0) = \sigma_Y^2$.

Problems with Solutions Provided at the End of the Text

1*. For the Exponential distribution : $dF = y_0 e^{-x/\sigma}\,dx,\ 0 \le x < \infty, \sigma > 0$

find the interquartile range, μ'_r and k_r. Show further that $\beta_1 = 4$ and $\beta_2 = 9$.

2*. If X is a variate with zero mean and cumulants k_r, show that the first three comulants λ_r of $Y = X^2$ are given by

$\lambda_1 = k_2,\ \lambda_2 = k_4 + 2k_2^2,\ \lambda_3 = k_6 + 2(5k_3^2 + 6k_2 k_4) + 8k_2^3$.

3*. Show that $M(t : X + Y) = M(t : X) \,.\, M(t : Y) \Rightarrow X$ and Y are uncorrelated.

4*. Let $X_1, X_2, \ldots, X_n$ be a random sample from a distribution with $\mathrm{E}\,|\,X^4\,| < \infty$. Write $\mu = \mathrm{E}(X)$ and $\mu_k = \mathrm{E}(X-\mu)^k,\ \overline{X} = (\Sigma X_j)/n$ and $S^2 = \Sigma(X_j - \overline{X})^2/(n-1)$. Prove :

$$\mathrm{Var}\,(S^2) = \frac{\mu_4 - 3\mu_2^2}{n} + \frac{2\mu_2^2}{n-1}, \qquad \mathrm{Cov}\,(\overline{X}, S^2) = \frac{\mu_3}{n}.$$

Exercise 8(b)

1. Write $m_K(t) = M^{(K)}(t)/M(t)\ \ K(t) = \ln M(t)$ [Cumulant generating function]

   Show that $m_k'(t) = m_{k+1}(t) - m_k(t)\,m_1(t)$

   Deduce relations for c.g.f. :

   $K'(t) = m_1(t),\ \ K''(t) = m_2(t) - [m_1(t)]^2$

   $K^{(3)}(t) = m_3(t) - 3m_2(t)\,m_1(t) + 2[m_1(t)]^3$

   $K^{(4)}(t) = m_4(t) - 4m_3(t)\,m_1(t) - 3m_2^2(t) + 12m_2(t)\,m_1^2(t) - 6m_1^4(t)$

2. Show that if $\mathrm{E}(X^r) = (r+1)!\,2^r$, then $k_r = (r-1)!\,(2)^{r+1}$.

3. For the distribution $f(x) = ce^{-cx}, x \ge 0, c > 0$, show $k_r = (r-1)!\,c^{-r}$.

4. Show that $k_r = 1.3.5..(2r-3)\ \mu^{2r-1}\lambda^{1-r}$. for the distribution

$$f(x) = \left(\frac{\lambda}{2\pi x^3}\right)^{1/2} \exp\left[\frac{-\lambda(x-\mu)^2}{2\mu^2 x}\right], x > 0, \lambda > 0, \mu > 0.$$

5. Show that $\dfrac{\partial \mu'_r}{\partial k_j} = \dbinom{r}{j}\mu'_{r-j},\quad \dfrac{\partial \mu_r}{\partial K_j} = \dbinom{r}{j}\mu_{r-j}\quad j = 1, 2, 3, \ldots r.\ (r \ge 2)$.

6. For the p.m.f. : $P(X=x)=a_x\theta^x/f(\theta), 0\le x<\infty\ \theta>0$ where $a_x\ge 0, f(\theta)=\Sigma a_x\theta^x$, prove that

$$\mu_r'=\sum_{j=1}^{\infty}\binom{r-1}{j-1}\mu_{r-j}'k_j;\ k_{r+1}=\theta\frac{dk_r}{d\theta},\ k_{r+1}=\theta\sum_{j=1}^{\infty}\binom{r-1}{j-1}\mu_{r-j}'\frac{dk_j}{d\theta}-\sum_{j=2}^{\infty}\binom{r-1}{j-2}\mu_{r+1-j}'k_j.$$

7. For the p.m.f. of Exercise (6), show that $k(r)=\theta^r D^r[\log f(\theta)]$. $[D=d/d\theta]$

8. Define a factorial cummulant $k_{(r)}$ as the coefficient of $t^r/r!$ in the expansion of the logarithm of the factorial m.g.f. Prove the following relations :

$$k_{(1)}=k_1,\ k_{(2)}=k_2-k_1,\ k_{(3)}=k_3-3k_2+2k_1,\ k_{(4)}=k_4-6k_3+11k_2-6k_1.$$

9. Prove that $\mu_r=\sum_{j=1}^{\infty}\binom{r-1}{j-1}k_j\mu_{r-j}$

10. Let $a_r=k_r/r!$. Then show that

$$\mu_r'=\begin{vmatrix} a_1 & -1 & 0 & 0 & \cdots & 0\\ 2a_2 & a_1 & -2 & 0 & \cdots & 0\\ 3a_3 & 2a_2 & a_1 & -3 & \cdots & 0\\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots\\ ra_r & (r-1)a_{r-1} & (r-2)a_{r-2} & \vdots & \cdots & a_1\end{vmatrix}$$

8-50. *Probability Generating Function*

The probability generating function (written p.g.f.) of an *integer-valued* variate X, written $G(t:X)$ or $G_X(t)$, is defined by the relation

$$G(t:X)=\mathsf{E}(t^X)=\Sigma_x P(X=x)\,t^X=\Sigma_x p(x)\,t^x. \qquad ...(1)$$

Note that the name p.g.f. is due to the fact that the coefficient of t^X in $\mathsf{E}(t^X)$ $P(X=x)$.

Uniqueness theorem. Let X and Y have p.g.f.'s : G_X and G_Y. Then

$$G_X(t)=G_Y(t)\ \forall\, t\Leftrightarrow P(X=k)=P(Y=k), 0\le k<\infty.$$

Proof. Since $\sum_{k=0}^{\infty}|p_k t^k|\le\sum_{k=0}^{\infty}p_k=1$, it follows that $G_Z(t)$ exists for all values of t satisfying $|t|\le 1$. Hence G_X and G_Y have radii of convergence at least 1 and therefore these have unique power series expansions about the origin

$$G_X(t)=\sum_{k=0}^{\infty}t^k\,p_k,\ G_Y(t)=\sum_{k=0}^{\infty}t^k q_k. \qquad [p_k=P(X-k), q_k=P(Y=k)]$$

If $G_X=G_Y$, then these power series have identical coefficients ($p_k=q_k$) and conversely, if $p_k=q_k$, we get $G_X=G_Y$.

Note. Physical meaning of p.g.f.

Let X be a non-negative integer valued variate with $P(X=k)=P_k$. Let t be a real number, $0\le t\le 1$. When $X=k$, perform k alternative trials each with probability of success t. Let A denote an event of *no failure*. Then, by Multi-Stage Rule

$$P(A)=\sum_{k=0}^{\infty}P(X=k)\,P(A\mid X=k)=\sum_{k=0}^{\infty}p_k t^k=G_X(t).$$

Relation between the P.G.F. and M.G.F.

(i) $M(t:X) = \mathsf{E}(e^{tX})$ (ii) $G(t:X) = (e^X)$

Thus (i) is obtainable from (ii) by changing t to e^t. Conversely, to obtain (ii) from (i), change e^t to t. So

$$M(t:X) = G(e^t:X); \quad G(t:X) = M(\ell n\ t:X)$$

Thus : $M(t:X) = (q+pe^t)^n$, then $G(t:X) = (q+pt)^n$.

Similarly $G(t:X) = e^{m(t-1)} \Rightarrow M(t:X) = e^{m(e^t-1)}$, etc.

8-51. *P.G.F.'s of Some Standard Distributions*

The computations for m.g.f. [§8-16] and p.g.f. are similar, instead of e^t we have to insert t.

1. ***Binomial distribution*** : $f(x) = \binom{n}{x} q^{n-x} p^x$, $x = 0, 1, \ldots, n$; $G(t) = (q+pt)^n$.

2. ***Poisson distribution*** : $f(x) = e^{-\lambda}\lambda^x / x!$, $x = 0, 1, 2, \ldots$ $G(t) = e^{\lambda(t-1)}$.

3. ***Geometric distributions*** : $f(x) = q^x p$, $x = 0, 1, 2, \ldots,$ $G(t) = p/(1-qt)$ $|qt| < 1$.

4. ***Pascal geometric distribution*** : $f(x) = q^{x-1}p$, $x = 1, 2, \ldots$ $G(t) = pt/(1-qt)$.

5. ***Neg-bin distribution*** : $f(x) = \binom{x+k-1}{k-1} p^k q^x$, $x = 0, 1, 2, \ldots$ $G(t) = [p/(1-qt)]^k$.

6. ***Pascal N-B distribution*** : $f(y) = \binom{y-1}{k-1} p^k q^{y-k}$, $y = k, (k+1), \ldots$ $G(t) = [pt/(1-qt)]^k$.

7. ***Multinomial distribution*** : $f(x_1, \ldots, x_k) = \binom{n}{x_1, \ldots, x_k} p_1^{x_1} \ldots p_k^{x_n}$, $G(t) = (p_1 t_1 + \ldots + p_k t_k)^n$.

8-52. *Effect of Linear Transformations on P.G.F.*

$$G(t:a+bX) = t^a\, G(t^b:X).$$

Proof. $G(t:a+bX) = \mathsf{E}(t^{a+bX}) = \mathsf{E}(t^a \,.\, t^{bX}) = t^a \mathsf{E}[(t^b)^X] = t^a G(t^b:X)$.

8-53. *Linear Property*

If X, Y are *independent* variates, then for constants a, b

$$G(t:aX+bY) = G(t^a:X)\, G(t^b:Y)$$

Proof. $G = (t:aX+bY) = \mathsf{E}[t^{aX+bY}] = \mathsf{E}[t^{aX} \,.\, t^{bY}] = \mathsf{E}[t^{aX}]\,\mathsf{E}[t^{bY}] = G(t^a:X)\, G(t^b Y)$

[$\because$ X and Y are indep.] In particular,

$$G(t:X+Y) = G(t:X)\, G(t:Y).$$

Generalization. If X_i, $i = 1, 2, \ldots, n$ are *independent* variates, then

$$G(t:X_1 + X_2 + \ldots + X_n) = G(t:X_1)\, G(t:X_2) \ldots G(t:X_n)$$

$$= [G(t:X_1)]^n, \text{ if } X_i\text{'s are i.i.d.}$$

The proof is immediate by mathematical induction.

8-54. Sums of Special Independent Random Variables

1. Let $X \sim$ bin (m, p) and $Y \sim$ bin (n, p) be *independent*. Then $X + Y \sim$ bin $(m + n, p)$.
2. Let $X \sim$ Pois (λ_1) and $Y \sim$ Pois (λ_2) be *independent*. Then $(X + Y) \sim$ Pois $(\lambda_1 + \lambda_2)$.
3. Let $X \sim NB(k_1, p)$ and $Y \sim NB(k_2, p)$ be *independent*. Then $(X + Y) \sim NB(k_1 + k_2, p)$.
4. Let $X \sim$ gem (p) and $Y \sim$ gem (p) be *independent*. Then $(X + Y) \sim NB(2, p)$.

Proof. In each case we use $G(t : X + Y) = G(t : X) \,.\, G(t : X)$ and appeal to uniqueness theorem.

1. $G(t:X) = (q + pt)^m, G(t:Y) = (q + pt)^n,\ G(t:X).G(t:Y) = (q + pt)^{m+n} = G(t:X+Y)$

Thus $(X+Y) \sim$ bin $(m+n, p)$.

2. $G(t:X) = e^{\lambda_1(t-1)}, G(t:Y) = e^{\lambda_2(t-1)}, G(t:X).G(t:Y) = e^{(\lambda_1+\lambda_2)(t-1)} = G(t:X+Y)$.

Thus $(X + Y) \sim$ Pois $(\lambda_1 + \lambda_2)$

3. $G(t:X) = [p/(1-qt)]^{k_1}, G(t:Y) = [p/(1-qt)]^{k_2}, G(t:X).G(t:Y) = [p/(1-qt)]^{k_1+k_2} = G(t:X+Y)$

Thus $(X + Y) \sim NB(k_1 + k_2, p)$.

4. $G(t:X) = p/(1-qt), G(t:Y) = p/(1-qt), G(t:X).G(t:Y) = [p/(1-qt)]^2 = G(t:X+Y)$.

Thus $(X + Y) \sim NB(2, p)$.

8-55. Moments via P.G.F.

Let X be an integer-valued random variable with a p.g.f. $G(t)$. Then

(a) $G(1) = 1$, (b) $\mu'_{(r)} = \mathsf{E}[(X^{(r)}] = G^r(1)$, (c) $\mathsf{E}(X) = G'_1$, (d) Var $(X) = G''(1) + G'(1) - [G'(1)]^2$

Proof. (a) By definition : $G(t) = \mathsf{E}(t^X)$ so that $G(1) = \mathsf{E}(1^X) = \mathsf{E}(1) = 1$.

$$G(t) = \mathsf{E}(t^X) = \Sigma_X t^x p(x). \qquad \text{...(i)}$$

We assume that the series in (i) is uniformly convergent, so that term by term differentiation is valid. Differentiating Eq. (i) r times w.r.t. 't' we get

$$G^r(t) = \sum_{x=r}^{\infty} x(x-1)(x-2)\ldots(x-r+1)\, t^{x-r} p(x)$$

where $G^r(t) = d^r G(t)/dt^r$. Putting $t = 1$, this gives

$$G^r(t) = \sum_{x=r}^{\infty} x(x-1)(x-2)\ldots(x-r+1)\, p(x) = \mathsf{E}[X^{(r)}] = \mu'_{(r)} \qquad \text{...(ii)}$$

(c) In particular, putting $r = 1$ in (ii), we get $G'(1) = \mu'_{(1)} = \mathsf{E}(X)$ ...(iii)

(c) Letting $r = 2$ in (ii), we get $G^2(t) = \mathsf{E}[X(X-1)] = \mu'_{(2)}$

$\therefore\ G''(1) = \mathsf{E}(X^2) - \mathsf{E}(X) \Rightarrow \mathsf{E}(X^2) = G''(1) + G'(1).$ [by (iii)]

$$\text{Var}(X) = \mathsf{E}(X^2) - \mathsf{E}^2(X) = G''(1) + G'(1) - [G'(1)]^2.$$

Example : The Dirichlet's p.g.f. has the form

$$G(t) = \Sigma(p_n / n^t),\ n \geq 1, \quad (G(0) = \Sigma p_n = 1) \qquad ...(1)$$

If X is a r.v. with $P(X = n) = p_n$, find Var(X) and E ($\ln X$)

Solution. Put $t = -1$ in (1) to get $G(-1) = \Sigma n . p_n = \mathrm{E}(X)$.

Put $t = -2$ in (1) to get $G(-2) = \Sigma n^2 p_n = \mathrm{E}(X^2)$

Thus Var $(X) = \mathrm{E}(X^2) - \mathrm{E}^2(X) = G(-2) - [G(-1)]^2$

Now differentiate (1) w.r.to 't' to get

$$G'(t) = \Sigma p_n \frac{d}{dt}(n^{-1})^t = \Sigma p_n (n^{-1})^t \ln (n^{-1})$$

Put $t = 0$, to get

$$G'(0) = \Sigma p_n (-1) \ln n = (-1) \Sigma p_n \ln n = (-1) \,|\, \mathrm{E}(\ln X)]$$

Thus $\mathrm{E}(\ln X) = -G'(0)$.

8-56. *Probability Mass Function via p.g.f.*

There is a unique one-to-one relation between the p.m.f. of a non-negative integer valued r.v. and its p.g.f.

Proof. Since $G(t) = \Sigma p_k t^k, 0 \leq k < \infty$, is convergent for all $|\ t\ | \leq 1$, term by term differentiation of the power series $G(t)$ is valid. Hence, for $|\ t\ | < 1$,

$$G'(t) = \sum_{k=1}^{\infty} k\ p_k\, t^{k-1}, G''(t) = \sum_{k=1}^{\infty} k^{(2)} t^{k-2}, ..., G^n(t) = \sum_{k=n}^{\infty} k^{(n)}\, t^{k-n}, ... [k^{(r)} = k(k-1) ... (k-r+1)].$$

Substituting $t = 0$, these give $G'(0) = p_1, G''(0) = 2!\, p_2, ..., G^n(0) = n!\, p_n$.

$$\therefore \qquad \boxed{G(t) = \sum_{n=0}^{\infty} \frac{G^n(0)}{n!} t^n \ ; \ p_k = \frac{1}{k!} G^k(0)}$$

It follows that the p.m.f. uniquely determines its p.g.f. and conversely the p.m.f. can be uniquely recovered from its p.g.f. This justifies the name to $G(t)$ as p.g.f.

8-57. *Product of Two Generating Functions*

$$G_1(t) . G_2(t) = (\Sigma p_i t^i)(\Sigma q_j t^j) = (p_0 + p_1 t + p_2 t^2 + ...)(q_0 + q_1 t + q_2 t^2 + ...)$$

$$= p_0 q_0 + (p_0 q_1 + p_1 q_0) t + (p_0 q_2 + p_1 q_1 + p_2 q_0) t^2 + ... + (p_0 q_n + p_1 q_{n-1} + ... + p_n q_0) t^n + ...$$

If we write $G(t) = G_1(t) . G_2(t)$ and let $G(t) = \Sigma r_n\, t^n$, it follows that

$$r_n = p_0 q_n + p_1 q_{n-1} + p_2 q_{n-2} + ... + p_n q_2 = \Sigma p_i q_{n-i},\ 0 \leq i \leq n. \qquad ...(1)$$

Whenever a sequence $\langle r_n \rangle$ is obtained from two sequences $\langle p_n \rangle$ and $\langle q_n \rangle$ by means of Eq. (1), we say that $\langle r_n \rangle$ is a *convolution* of $\langle p_n \rangle$ and $\langle q_n \rangle$, written $r = p * q = q * p$.

Conversely, $r = p * q \Rightarrow G(t) = G_1(t) . G_2(t)$.

Cor. 1. Let X and Y be indep. integer-valued variates with probabilities $\langle p_n \rangle$ and $\langle q_n \rangle$ and let $Z = X + Y$. Then $\{Z = n\} = \{(0, n) \uplus (1, n-1) \uplus (2, n-2) \uplus ... \uplus (n, 0)\}$,

$$P(Z=n)=p_0q_n+p_1q_{n-1}+\ldots+p_nq_0=\Sigma p_iq_{n-i}.$$

Thus $r = p * q$, so that $G(t:Z) = G(t:X)\,G(t:Y)$.

***Cor.* 2.** The generating function of $P(X \le n)$ is $H(t) = G(t)/(1-t)$. $[n \ge 0]$

Proof. Let $h_n = P(X \le n) = p_0 + p_1 + p_2 + \ldots + p_n = 1\,.\,p_0 + 1\,.\,p_1 + 1\,.\,p_2 + \ldots + 1\,.\,p_n$

Thus, $h = 1 * p$ and $H(t) = (1-t)^{-1}\,.\,G(t) = G(t)/(1-t)$ becuase the generating function for the sequence <1> is $1/(1-t)$.

$$[G(t)=\mathrm{E}(t^X)=1.1+1.t+1.t^2+1.t^3+\ldots=1+t+t^2+\ldots=(1-t)^{-1}]$$

8-58. *Convolution (Statistical Sum) Formula*

Let X and Y be *independent* integer-valued variates and $Z = X + Y$. Suppose

$$P(X=k)=p_k,\ P(Y=j)=q_j,\ P(Z=k)=r_k.\quad j,k=0,1,2,3,\ldots$$

Then $P(X+Y=k)=\Sigma P(X=i)\,P(Y=k-i)$, or $r_k=\Sigma p_iq_{k-i},\ 0\le i<k$.

Proof. Here $G(t:X+Y)=G(t:X).G(t:Y)\Rightarrow \Sigma r_kt^k=(\Sigma p_it^i)(\Sigma q_jt^j)\Rightarrow\ \Sigma r_k\,t^k=\Sigma\Sigma p_iq_jt^{i+j}$.

Comparing the coefficient of t^k on both sides ($i + j = k$) yields

$$r_k=\Sigma_ip_iq_{k-i},\ \text{i.e.}\ P(X+Y=k)=\Sigma P(X=i)\,P(Y=k-i).$$

8-60. *Some Tail-ends Generating Functions*

To Find the generating functions of :

(a) $P(X > n)$, (b) $P(X \ge n)$, (c) $P(X < n)$, (d) $P(X \le n)$, (e) $P(x = 2n)$.

Proof. Recall : $G(t)=\Sigma p_nt^n, n=0,1,2,3,\ldots\ [p_n=P(X=n)]$

(a) $P(X>n)-P(X>n+1)=P(X=n+1)\Rightarrow f_n-f_{n+1}=p_{n+1}$ $[f_n = P(X > n)]$

$$\therefore\quad t^nf_n-(1/t)t^{t+1}f_{n+1}=(1/t)t^{n+1}p_{n+1}$$

So $\displaystyle\sum_{n=0}^{\infty}t^nf^n-\frac{1}{t}\sum_{n=0}^{\infty}t^{n+1}f_{n+1}=\frac{1}{t}\sum_{n=0}^{\infty}t^{n+1}p_{n+1}$ [Put $H(t)=\displaystyle\sum_{n=0}^{\infty}f_nt^n$] ...(i)

Now, $\displaystyle\sum_{n=0}^{\infty}t^{n+1}f_{n+1}=\sum_{n=0}^{\infty}f_nt^n-f_0=H(t)-f_0;\ \sum_{n=0}^{\infty}p_{n+1}t^{n+1}=\sum_{n=0}^{\infty}p_nt^n-p_0=G(t)-p_0$

Further, $f_0=P(X>0)=1-P(X=0)=1-p_0$. Making substitutions into (i) we get

$$H(t)-t^{-1}[H(t)-(1-p_0)]=t^{-1}[G(t)-p_0]\Rightarrow H(t)=[1-G(t)]/(1-t) \quad \text{...(a)}$$

Formula (a) gives the generating function $H(t)$ of the *tail probabilities.*

(b) $P(X\ge n)=P(X=n)+P(X>n)$, i.e. $g_n=p_n+f_n$ [with obvious notation]

$$\therefore\quad \Sigma g_n\,t^n=\Sigma p_nt^n+\Sigma f_nt^n \qquad (n = 0, 1, 2, \ldots)$$

or $H_1(t)=G(t)+H(t)=G(t)+[1-G(t)]/[1-t],=[1-tG(t)]/(1-t).$ [by (a)] ...(b)

(c) From relation : $P(X < n + 1) - P(X < n) = P(X = n)$, we get

$$\phi_{n+1}-\phi_n=p_n,\ [\phi_n=P(X<n)] \qquad (n = 0, 1, 2, 3, \ldots, \infty)$$

$\therefore \quad t^{-1}(\Sigma t^{n+1}\phi_{n+1}) - (\Sigma t^n \phi_n) = (\Sigma t^n p_n) \Rightarrow t^{-1}[H_2(t) - \phi_0] - H_2(t) = G(t)$

As $\phi_0 = P(X < 0) = 0$, the above result reads

$$H_2(t) = tG(t)/(1-t) \qquad [H_2(t) = \Sigma P(X<n)t^n] \qquad \ldots(c)$$

Aliter. $P(X<n) + P(X=n) + P(X>n) = 1$ yields

$$\Sigma t^n\, P(X<n) + \Sigma t^n\, p_n + \Sigma t^n P(X>n) = \Sigma t^n, \qquad [0 \le n \le \infty]$$

$$\therefore \quad H_2(t) + G(t) + H(t) = (1-t)^{-1} \Rightarrow H_2(t) = tG(t)/(1-t). \qquad \text{[by (a)]}$$

(d) $\Sigma t^n\, P(X \le n) = \Sigma t^n[1 - P(X>n)] = \Sigma t^n - \Sigma t^n P(X>n)$

$$= \frac{1}{1-t} - \frac{1-G(t)}{1-t} = \frac{G(t)}{1-t}.$$

(e) Let $P(X = 2n) = p_{2n}$, then

$$H_4(t) = \Sigma p_{2n} t^n = \Sigma p_{2n}(\sqrt{t})^{2n} = \Sigma p_{2n} s^{2n}. \quad (s = \sqrt{t}), n = 0, 1, 2, \ldots$$

$$= \tfrac{1}{2}\sum_{x=0}^{\infty}[p_k s^k + p_k(-s)^k] = \tfrac{1}{2}[G(s) + G(-s)] = \tfrac{1}{2}[G(\sqrt{t}) + G(-\sqrt{t})]$$

$$\boxed{\Sigma t^n P(X>n) = \frac{1-G(t)}{1-t};\ \Sigma t^n\, P(X \ge n) = \frac{1-tG(t)}{1-t};\ \Sigma t^n P(X<n) = \frac{tG(t)}{1-t};\ \Sigma t^n P(X \le n) = \frac{G(t)}{1-t}}$$

8-61. *Worked-out Problems*

Example 1. Find the p.g.f. and m.g.f. of the distribution $f(x) = \frac{1}{2}pq^{|x|-1}$, $x = \pm 1, \pm 2, \pm 3, \ldots, 0 < p, q < 1, p + q = 1$.

Solution.

$$G(t) = \mathsf{E}(t^X) = \sum_x \tfrac{1}{2}pq^{|x|-1}.t^x$$

$$= \tfrac{1}{2}(p/q)\sum_x t^x.q^{|x|}$$

$$= (p/2q)\{q(t+t^{-1}) + q^2(t^2+t^{-2}) + q^3(t^3+t^{-3}) + \ldots\}$$

$$= (p/2q)\{\Sigma(qt)^r + \Sigma(q/t)^r\}, \quad 1 \le n < \infty$$

$$= \frac{p}{2q}\left\{\frac{qt}{1-qt} + \frac{q/t}{1-(q/t)}\right\}, \quad |qt| < 1, |q| < t.$$

$$= \frac{p}{2}\left(\frac{t}{1-qt} + \frac{1}{t-q}\right), \quad q < |t| < 1/q$$

For m.g.f., we just replace t by e^t. Thus

$$M(t) = \frac{p}{2}\left(\frac{e^t}{1-qe^t} + \frac{1}{e^t - q}\right), \ \ln q < t < -\ln q.$$

***Example* 2.** Show that (a) $G(t) = 2/(1 + t)$ is *not* a p.g.f.
(b) If $G(t)$ is a p.g.f., is $[2 - G(t)]^{-1}$ a p.g.f. ?

Solution. We recall that if $G(t) = p_0 + p_1 t + p_2 t^2 + \ldots$ is a power series that converges when $|t| \le 1$, then $G(t)$ is a p.g.f. iff $p_r \ge 0$ for all r and $G(1) = 1$. Here

(a) Here $G(t) = 2(1 + t)^{-1} = 2\Sigma\,(-t)^r \;\Rightarrow\; p_{2k+1} < 0.$

Since some coefficients are negative, $G(t)$ is not a p.g.f.

(b) $$H(t) = \tfrac{1}{2}[1 - \tfrac{1}{2}G(t)]^{-1} = \sum_{r=0}^{\infty}\left(\tfrac{1}{2}\right)^{r+1}[G(t)]^r.$$

Now, coefficient of t^n in each of $G(t)$, $[G(t)]^2$, $[G(t)]^3$, ... is non-negative because these are each p.g.f.'s. If we express (1) as a polynomial in t, then

$$H(t) = p_0 + p_1 t + p_2 t^2 + \ldots$$

each $p_j \ge 0$ and $p_0 + p_1 + p_2 + \ldots = [2 - G(1)^{-1} = (2 - 1)^{-1} = 1$. Thus, $H(t)$ is indeed a p.g.f.

***Example* 3.** If $G(t)$ is the p.g.f. of a r.v. X such that $P(X = n) = p_n$, obtain the p.g.f. of $P(X = 2n)$. If G_1 and G_2 are p.g.f.'s then $G(t) = G_2\,[G_1\,(t)]$ is also a p.g.f.

Solution. Since G_1 and G_2 are p.g.f.'s we have

$$G(t) = G_2[G_1(t)] = p_0 + p_1\,G_1(t) + p_2\,[G_1(t)]^2 + \ldots + p_r[G_1(t)]^r + \ldots \qquad \ldots(1)$$

$$\therefore\; G(1) = p_0 + p_1 G_1(1) + p_2[G_1(1)]^2 + \ldots = p_0 + p_1 + p_2 + \ldots = 1\,[G_1(1) = 1, \text{as } G_1 \text{ is p.g.f.}]$$

We rewrite (1) as

$$G(t) = p_0 + p_1\,[P_0 + P_1 t + P_2 t^2 + \ldots] + p_2\,[P_0 + P_1 t + P_2 t^2 + \ldots]^2 + \ldots$$

This shows that each coefficient of t^r is non-negative. Since the power series (1) converges when $|t| \le 1$, $G(1) = 1$, and coefficient of t^r for all r, are non-negative, we conclude that $G(t)$ is indeed a p.g.f.

***Example* 4.** Let $\{X_k\}$ be mutually independent variates each assuming the values 0, 1, 2, ..., $a - 1$ with probabilities $1/a$. If $S_n = X_1 + X_2 + \ldots + X_n$, show that the p.g.f. of S_n is

$$G(t) = [(1 - t^a)/(1 - t)\,a]^n.$$

$$\text{Deduce : } P(S_n = j) = \frac{1}{a^n}\sum_{k=0}^{\infty}\binom{n}{k}\binom{-n}{j - ak}(-1)^{k+j+ak}.$$

Solution. Here, $G(t : X_i) = \mathrm{E}(t^{X_i}) = \displaystyle\sum_{k=0}^{a-1}\left(\frac{1}{a}\right)t^k = \frac{1 - t^a}{a(1 - t)}$;

Since X_i are i.i.d. variates, we get

$$G(t : S_n) = [G(t : X_i)]^n = [(1 - t^a)/a(1 - t)]^n = (1 - t^a)^n\,(1 - t)^{-n}/a^n$$

$$= \frac{1}{a^n}\sum_{k=0}^{n}\binom{n}{k}(-t^a)^k \cdot \sum_{r=0}^{\infty}\binom{-n}{r}(-1)^r = \frac{1}{a^n}\sum_{k=0}^{n}\sum_{r=0}^{\infty}\binom{n}{k}\binom{-n}{r}(-1)^{k+r}\,t^{ak+r}. \qquad \ldots(1)$$

Now $p(S_n = j)$ = Coeff. of t^j in $G(t : S_n)$; hence putting $r + ak = j$ in (1) we get this Coeff. as

$$P(S_n = j) = \frac{1}{a^n} \sum_{k=0}^{n} \binom{n}{k} \binom{-n}{j-ak} (-1)^{k+j-ak} \qquad ...(2)$$

Since $\qquad (-1)^{2ka} = 1 \Rightarrow (-1)^{ka} = (-1)^{-ka}$; also $\binom{n}{k} = 0,$, if $k > n$; So (2) yields

$$P(S_n = j) = \frac{1}{a^n} \sum_{k=0}^{\infty} \binom{n}{k} \binom{-n}{j-ak} (-1)^{k+ak+j}.$$

***Example* 5.** Let $\{X_k\}$ be a sequence of independent variates with a common distribution given by a generating function $F(t)$ and let $S_N = X_1 + X_2 + ... + X_N$. If N is itself a variate independent of X_i and has a p.g.f. $G(t)$, show that p.g.f. of S_N is given by $GF(t)$. Hence, show that

$$\mathrm{E}(S_N) = \mathrm{E}(N)\mathrm{E}(X); \mathrm{Var} = (S_N) = \mathrm{E}(N) \mathrm{Var}(X) + \mathrm{E}^2(X) \mathrm{Var}(N).$$

Solution. Since X_i are i.i.d. variates, we have

$$g(t : S_n) = g(t : X_1 + ... + X_n) = [g(t : X_1)^2] = [F(t)]^n = (\theta)^n . \; [\theta = F(t)]$$

$\therefore \; g(t : S_N) = \mathrm{E}(t^{S_N}) = \mathrm{E}\{\mathrm{E}(t^{S_N} / N)\} = \mathrm{E}(\theta^N). \qquad$ [p.g.f. of N] [by Double E-Rule]

By hypothesis, $G(t) = \mathrm{E}(t^N)$, it follows from above that

$$g(t : S_N) = G(\theta) = G(F(t)) = H(t), \text{ say}$$

$$H(1) = G(F(1)) = G(1) = 1. \qquad [F(t) = \mathrm{E}(t^X)]$$

$$H'(t) = G'(F) . F'(t), H''(t) = G''(F) . [F'(t)]^2 + G'(F) . F''(t).$$

$$\therefore \qquad \mathrm{E}(S_N) = H'(1) = G'(1) . F'(1) = E(F) . E(X)$$

As, $\qquad H''(1) = G''(1) [F'(1)]^2 + G'(1) F''(1) = \mathrm{E}[N(N-1)] \mu_X^2 + \mu_N . \mathrm{E}[X(X-1)]$

$$\therefore \qquad \mathrm{Var}(S_N) = H''(1) + H'(1) - [H'(1)]^2$$

$$= \{[\mathrm{E}(N^2) - \mu_N] \mu_X^2 + [\mathrm{E}(X^2) - \mu_X] \mu_N + \mu_N . \mu_X - \mu_N^2 . \mu_X^2$$

$$= [\mathrm{E}(N^2) - \mu_N^2] \mu_X^2 + [\mathrm{E}(X^2) - \mu_X^2] \mu_N = \sigma_X^2 \mu_N + \sigma_N^2 \mu_X^2.$$

The mean and variance of S_N is also obtained in Wald's thoerem § 5-54.

Note. $\qquad M(t : S_n) = [M(t : X)]^n$ since X_i are i.i.d.

$$\therefore \qquad M(t : S_N) = \mathrm{E}(e^{tS_N}) = \mathrm{E}[\mathrm{E}(e^{tS_N} \mid N)] = \mathrm{E}[M_X(t)]^N \, \mathrm{E}(\theta^N) = G(\theta : N).$$

***Example* 6.** A hen lays X eggs where X is Pois (λ). Let Y be the number of eggs which hatch independentily with probability p and Z be the those which donot hatch. (no chicks). Show that Y and Z are independent. Further show Corr $(X, Y) = \sqrt{p}$.

Solution. When conditioned $X = n$, then $Y \mid X$ is bin $(.\imath, p)$ so that

$$G(t : Y \mid X) = (q + pt)^n \qquad \text{[p.g.f. of bin } (n, p)] \qquad ...(i)$$

Now we consider joint p.g.f. of Y and Z :

$$\begin{aligned}
\mathsf{E}(t_1^Y . t_2^Z) &= \mathsf{E}(t_1^Y . t_2^{X-Y}) = \mathsf{E}[t_2^X . (t_1 / t_2)^Y] \qquad [\text{Let } t_1/t_2 = \theta] \\
&= \mathsf{E}\{\mathsf{E}(t_2^X . \theta^Y \mid X = x)\}, \qquad [\text{by Double-E Rule}] \\
&= \mathsf{E}\{t_2^X . \mathsf{E}(\theta^Y \mid X = x)\} \qquad [Y \mid x \sim \text{bin } (n, p) \text{ and pullout property}] \\
&= \mathsf{E}\{t_2^X . (q + p\theta)^X\} = \mathsf{E}\{(qt_2 + pt_1)^X\}, \qquad [t_3 = qt_2 + pt_1] \\
&= \mathsf{E}(t_3^X) = e^{\lambda(t_3 - 1)} \qquad [X \sim \text{Pois } (\lambda)] \\
&= e^{\lambda(qt_2 + pt_1) - \lambda(p+q)} = e^{\lambda p(t_1 - 1)} . e^{\lambda q(t_2 - 1)}
\end{aligned}$$

This shows that $Y \sim$ pois (λp) and $Z \sim$ Pois (λq) are indep.

Thus $\mathsf{E}(Y) = \lambda p = \text{Var}(Y)$ Now

$$\begin{aligned}
\mathsf{E}(XY) &= \mathsf{E}\{\mathsf{E}(XY \mid X = x)\} = \mathsf{E}\{X\, \mathsf{E}(Y \mid x)\}, \; [Y \mid X \sim \text{bin } (n, p)] \\
&= \mathsf{E}\{X . pX\} = p\mathsf{E}\{X^2\} = p\{\text{Var}(X) + \mathsf{E}^2(X)\} \\
&= p(\lambda + \lambda^2) \\
\text{Cov } (X, Y) &= \mathsf{E}(XY) - \mathsf{E}(X)\mathsf{E}(Y) = \lambda p + \lambda(\lambda p) - \lambda(\lambda p) = \lambda p \\
\text{Corr } (X, Y) &= \frac{\sigma_{XY}}{\sigma_X \sigma_Y} = \frac{\lambda p}{\sqrt{\lambda p \lambda}} = \sqrt{p}.
\end{aligned}$$

Problems with Solutions Provided at the End of the Text

1*. Let $X \sim$ Expo (λ) and Y be a discrete r.v. defined through X by

$$P\{Y = \tfrac{1}{2}(2n+1)h\} = P\{nh \le X \le (n+1)h\}, \; h > 0, \; n = 0, 1, 2, 3, \ldots$$

Find p.g.f. of Y and deduce that $\mathsf{E}(Y) > \mathsf{E}(X)$ but Var $(Y) <$ Var (X).

2*. Lottery tickets bear numbers from 000000 to 999999. Find the chance that a ticket bears a number whose sum of the first three digits equals the sum of the last three digits.

3*. A die is thrown n times. Let S be the total number of points. Show that

$$P(S = n + 5) = \binom{n+4}{5}\left(\frac{1}{6}\right)^n ; \; P(S = n + 4) = \binom{n+3}{4}\left(\frac{1}{6}\right)^n .$$

4*. There is a complete set of values (events) $x_0, x_1, \ldots, x_k, \ldots$ with probabilities $p_0, p_1, \ldots, p_k, \ldots$ which vary with time. These are related by the differential equation

$$dp_k/dt = \lambda(p_{k-1} - p_k) \qquad \ldots(1)$$

Find the value of p_k.

5*. A variate X assumes the values $x_1, x_2, \ldots$ with probabilities $p_1, p_2, \ldots$ Show that

$$P_k = (k!)^{-1} \sum_{j=0}^{\infty} p_j e^{-x_j} (x_j)^k \qquad [x_i > 0, \quad \Sigma p_i = 1] \qquad \ldots(1)$$

is a probability distribution of some variate Y. Find its p.g.f. and prove that $\mathsf{E}(Y) = \mathsf{E}(X)$ and Var $(Y) = \mathsf{E}(X) +$ Var (X).

6*. Let $X \sim$ geom (p) and $Y \sim$ geom (p') be independently distributed. Find the p.m.f. of $Z = X + Y$.

8-70. *Bivariate Probability Generating Function*

Let X and Y be non-negative integer-valued variates. The joint p.g.f. of X and Y is a function $G_{X,Y}$ of two variables t_1, t_2 defined by

$$G_{X,Y}(t_1, t_2) = \mathsf{E}(t_1^X . t_1^Y) = \sum_{x,y} t_1^x . t_2^y. \qquad \ldots(1)$$

The probabilities $f(x, y) \equiv P(X = x, Y = y)$ may be generated by expanding G as a power series in (t_1, t_2) and reading the coefficients.

The marginal p.g.f. of X and that of Y are obtained at once from (1)

$$G_{X,Y}(t_1, 1) = \mathsf{E}(t_1^X) = G_X(t_1); \; G_{X,Y}(1, t_2) = \mathsf{E}(t_2^Y) = G_Y(t_2). \qquad \ldots(2)$$

Moments. Assuming term by term differentiation of the power series (1), which is equivalent to interchange of differentiation operator and expectation operator E, we get

$$G^{(r,s)}(t_1, t_2) = \frac{\partial^{r+s}}{\partial t_1^r \, \partial t_2^s} G(t_1, t_2) = \mathsf{E}\{X^{(r)} Y^{(s)} t_1^{X-r} t_2^{Y-s}]$$

$$\therefore \qquad G^{(r,s)}(1, 1) = \mathsf{E}[X^{(s)} Y^{(s)}]. \; [t_1 = t_2 = 1] \qquad \ldots(3)$$

This gives the *descending* factorial moments of X, Y. An alternate derivation is through Power Series about (0, 0)

$$W_{X,Y}(t_1, t_2) = \mathsf{E}(1 + t_1)^X (1 + t_2)^Y = \sum_{i,j} t_1^i t_2^j \, \mathsf{E}(X^{(i)} Y^{(j)}) / (i! \; j!)$$

Thus, $\quad \mathsf{E}(X^{(i)} Y^{(j)}) = \text{Coeff. of } t_1^i t_2^j / i! \; j! \text{ in } W(t_1, t_2).$

***Factorization Theorem for indep.* variates.** Variates X and Y are independent iff

$$G_{X,Y}(t_1, t_2) = G_X(t_1) . G_Y(t_2)$$

Multivariate p.g.f. $G(t_1, t_2, \ldots, t_k) = \mathsf{E}(t_1^{X_1} . t_2^{X_2} \ldots t_k^{X_k}) = \sum_x t_1^{X_1} \ldots t_k^{X_k} f(x_1, \ldots, x_k)$.

When $|t_i| \le 1$, $\forall \; i \; (1 \le i \le k)$; then $G(t_1, 1, 1, \ldots 1) = G_{X_1}(t_1)$, etc.

8.71. *Worked-out Problems*

***Example* 1.** The r.vs. X and Y have joint p.m.f. :

$f(a, 0) = f(0, a) = f(-a, 0) = f(0, -a) = \frac{1}{4}$

where $f(x, y) = P(X = x, Y = y)$. Show that $X + Y$ and $X - Y$ are indep.

Solution. It pays to use joint p.g.f. Let $U = X + Y$, $V = X - Y$. Then

$$\begin{aligned}
G(t_1, t_2) &= \mathsf{E}\{t_1^{X+Y} . t_2^{X-Y}\} = \mathsf{E}\{(t_1 t_2)^X (t_1 t_2^{-1})^Y\} \\
&= \mathsf{E}\{t_3^X . t_4^Y\}, \quad t_3 = t_1 t_2,\ t_4 = t_1 t_2^{-1}. \\
&= \tfrac{1}{4}\{t_3^a + t_4^a + t_3^{-a} + t_4^{-a}\} \\
&= \tfrac{1}{4}\{t_1^a t_2^a + t_1^a t_2^{-a} + t_1^{-a} t_2^{-a} + t_1^{-a} t_2^a\} \\
&= \tfrac{1}{4}\{(t_1^a + t_1^{-a})(t_2^a + t_2^{-a})\} \\
&= G_U(t_1) . G_V(t_2)
\end{aligned}$$

$x \downarrow$ \ $y \rightarrow$	$-a$	0	a	$f_1(x)$
$-a$	0	1/4	0	1/4
0	1/4	0	1/4	1/2
a	0	1/4	0	1/4
$f_2(y)$	1/4	1/2	1/4	1

This shows that U and V are independent.

Note that X and Y are not independent.

$$0 = f(0, 0) \neq f_X(0) . f_Y(0) = \tfrac{1}{4}$$

Domain set of U as well as of V is

$$\{-2a, -a, 0, a, 2a\}$$

***Example* 2.** The p.g.f. of three discrete variates X, Y, Z is

$$G(t_1, t_2, t_3) = \mathsf{E}(t_1^X t_2^Y t_3^Z) = \frac{1}{2^n}\left[1 + \frac{t_1}{2^n}\left(1 + \frac{t_2(1+t_3)^n}{2^n}\right)^n\right]^n \qquad \ldots(1)$$

Using (1), show that

(a) $P(Y = 1) = \dfrac{n^2(2^n + 1)^{n-1}}{2^{n(n+1)}}$, (b) $P(X = 1 \mid Y = 1) = \dfrac{2^{n(n-1)}}{(2^n + 1)^{n-1}}$,

(c) $P(Z = 0) = \dfrac{[2^{n(n+1)} + (2^n + 1)^n]^n}{2^{n(n^2+n+1)}}$, (d) $\mathsf{E}(X + Y + Z) = \dfrac{n(n^2 + 2n + 4)}{2^3}$.

Solution. (a) $G(t_2) = G_Y(t) = 2^{-n}[1 + 2^{-n}(1+t)^n]^n$. $[t_1 = t_3 = 1 \text{ in } (1)]$

$$G_Y'(t) = n^2 . 2^{-2n}(1+t)^{n-1}[1 + 2^{-n}(1+t)^n]^{n-1}.$$

$$P(Y = 1) = G_Y'(0) = \frac{n^2}{2^{2n}}\left(1 + \frac{1}{2^n}\right)^{n-1} = \frac{n^2(2^n+1)^{n-1}}{2^{n(n+1)}}.$$

(b) $$G(t_1, t_2) = \frac{1}{2^n}\left[1 + \frac{t_1}{2^n}(1+t_2)^n\right]^n = \frac{1}{2^n}T^n, \text{ (say)}, \qquad [t_3 = 1 \text{ in } G(t_1, t_2, t_3)]$$

$$\frac{\partial G(t_1, t_2)}{\partial t_1} = \frac{n}{2^{2n}}(1+t_2)^n T^{n-1}; \quad \frac{\partial^2 G(t_1, t_2)}{\partial t_1 \partial t_2} = \frac{n^2(1+t_2)^{n-1}}{2^{2n}}T^{n-1} + \frac{n(n-1)}{2^{3n}}(1+t_2)^{2n-1} t_1 T^{n-2}$$

$$P(X = 1, Y = 1) = \partial^2 G(0, 0)/\partial t_1 \partial t_2 = n^2/2^{2n}.$$

$$\therefore\ P(X = 1 \mid Y = 1) = P(X = 1, Y = 1)/P(Y = 1) = 2^{n(n-1)}/(2^n + 1)^{n-1}.$$

(c) $G_Z(t) = \dfrac{1}{2^n}\left[1+\dfrac{1}{2^n}\left(1+\dfrac{(1+t)^n}{2^n}\right)^n\right]^n = \mathsf{E}(t^Z) = \Sigma_z t^z P(Z=z)$ $\qquad [t_1 = t_2 = 1$ in $G(t)]$

$$P(Z=0) = \text{Constant term in } G_Z(t) = \frac{1}{2^n}\left[1+\frac{(1+2^n)^t}{2^{n^2+n}}\right] = \frac{[2^{n^2+n}+(1+2^n)^n]^n}{2^{n(n^2+n+1)}}.$$

(d) Put $t_1 = t_2 = t_3\, t$; $X + Y + Z = S$; Then

$$G_S(t) = \mathsf{E}(t^S) = \frac{1}{2^n}\left[1+\frac{t}{2^n}\left(1+\frac{t(1+t)^n}{2^n}\right)^n\right]^n = \frac{1}{2^n}\theta^n, \text{ (say).}$$

$$G_S'(t) = \frac{n}{2^n}\theta^{n-1}.\left\{\frac{1}{2^n}\left(1+\frac{t(1+t)^n}{2^n}\right)^n + \frac{nt}{2^n}\left(1+\frac{t(1+t)^n}{2^n}\right)^{n-1}\left[\frac{(1+t)^n+nt(1+t)^{n-1}}{2^n}\right]\right\}$$

Put $t = 1$, to get

$$\mathsf{E}(S) = G_S'(t)\frac{n}{2^n}.2^{n-1}\left\{1+\frac{n}{2}\left[1+\frac{n}{2}\right]\right\} = \frac{n(n^2+2n+4)}{8}.$$

***Example* 3.** Suppose (X, Y) have joint p.m.f.

$$f(x, y) = \frac{e^{-(a+b)}a^x b^{y-x}}{x!(y-x)!},\ x = 0, 1, 2, \ldots, y = x,\ x+1, \ldots$$

Show that Corr $(X, Y) = [a/(a+b)]^{1/2}$ and obtain the distribution of $Z = Y - X$.

Solution. In the following summation, we use $z = x - y$.

$$G(t_1, t_2) = \mathsf{E}(t_1^X . t_2^Y) = \sum_{x=0}^{\infty}\sum_{y=x}^{\infty}\frac{e^{-(a+b)}a^x b^{y-x}}{x!(y-x)!}t_1^x . t_2^y = e^{-(a+b)}\sum_{x=0}^{\infty}\frac{(at_1t_2)^x}{x!}\sum_{X=0}^{\infty}\frac{(bt_2)^z}{z!}$$

$$= e^{-(a+b)}.e^{at_1t_2}.e^{bt_2} \qquad \ldots(1)$$

$$G_1(t) = G(t, 1) = e^{a(t-1)} \Rightarrow X \sim \text{Pois}(a), \text{ so } \mathsf{E}(X) = \text{Var}(X) = a.$$

$$G_2(t) = G(1, t) = e^{(a+b)(t-1)} \Rightarrow Y \sim \text{Pois}(a+b), \text{ so } \mathsf{E}(Y) = \text{Var}(Y) = a+b.$$

Put $t_1\, t_2 = 1$ in (1) to obtain

$$G(t^{-1}, t) = \mathsf{E}(t^{Y-X}) = e^{b(t-1)} \Rightarrow (Y-X) \sim \text{Pois (b). So Var}(Y-X) = b.$$

From Var $(X - Y) = \sigma_X^2 + \sigma_Y^2 - 2\sigma_X\sigma_Y\rho_{XY}$, we at once get

$$b = a + (a+b) - 2\rho\sqrt{a(a+b)} \Rightarrow \rho = [a/(a+b)]^{1/2}$$

***Example* 4.** Find the joint p.g.f. of the bivariate p.m.f.

$$f(x, y) = k\binom{x+y-1}{x}p^x Q^y,\quad x \ge 0,\ y \ge 1,\ 1 > 1-p > Q > 0. \quad [P+Q=1, p+q=1]$$

Determine ρ_{XY} also.

Solution. The constant k is determined by Normality $G(1, 1) = 1$. Now

$$G(t_1, t_2) = \mathsf{E}(t_1^X . t_2^Y) = k \sum_{y=1}^{\infty} \sum_{x=0}^{\infty} \binom{x+y-1}{x} (pt_1)^x . (Qt_2)^y$$

$$= k \sum_{y=1}^{\infty} (Qt_2)^y \left\{ \sum_{x=0}^{\infty} \binom{x+y-1}{x} (pt_1)^x \right\}. \qquad \left[\text{Apply } (1-T)^{-n} = \sum_{r=0}^{\infty} \binom{n+r-1}{r} T^r \right]$$

$$= k \sum_{y=1}^{\infty} (Qt_2)^y . (1-pt_1)^{-y} = k \sum_{y=1}^{\infty} \left(\frac{Qt_2}{1-pt_1} \right)^y \qquad \text{[Apply G.P. Sum]}$$

$$= k \left\{ \frac{Qt_2/(1-pt_1)}{1-[Qt_2/(1-pt_1)]} \right\} = \frac{kQt_2}{1-pt_1-Qt_2}.$$

By $G(1, 1) = 1$, we get $k = (q - Q)/Q$, so that

$$G(t_1, t_2) = (q - Q)\, t_2/(1 - pt_1 - Qt_2) \qquad \ldots(1)$$

$$G_X(t_1) = G(t_1, 1) = \frac{(q-Q)}{(P-pt_1)} = \frac{(q-Q)/P}{1-(p/P)t_1} \qquad (P + Q = 1) \qquad \ldots(2)$$

Recall that for geom (p), $G(t) = p/(1 - qt)$; it follows from (2) that $X \sim$ geom $[1 - (p/P)]$.

$$G_Y(t_2) = G(1, t_2) = \frac{(q-Q)t_2}{(q-Qt_2)} = \frac{\{1-(Q/q)\}t_2}{[1-(Q/q)t_2]} \qquad \ldots(3)$$

Recall that for gem (p), $G(t) = pt/(1 - qt)$; it follows from (2) that $Y \sim$ gem $[1 - (Q/q)]$. To determine ρ_{XY}, we need Cov (X, Y), so we find $\mathsf{E}(XY)$.

$$\frac{\partial^2 G(t_1, t_2)}{\partial t_1\, \partial t_2} = \frac{2pQ(q-Q)t_2}{(1-pt_1-Qt_2)^3} + \frac{p(q-Q)}{(1-pt_1-Qt_2)^2}$$

$$\mathsf{E}(XY) = \frac{\partial^2 G(1,1)}{\partial t_1\, \partial t_2} = \frac{p(q+Q)}{(q-Q)^2} = \frac{pQ}{(q-Q)^2} + \frac{pq}{(q-Q)^2} \qquad \ldots(4)$$

Now, the known results give

$$\mathsf{E}(X) = \frac{p/P}{1-(p/P)}, \ \mathrm{Var}(X) = \frac{p/P}{[1-(p/P)]^2} = \frac{pP}{(P-p)^2}$$

$$\mathsf{E}(Y) = \frac{1}{1-(Q/q)}, \ \mathrm{Var}(Y) = \frac{Q/q}{[1-(Q/q)]^2} = \frac{Qq}{(q-Q)^2}$$

$$\mathrm{Cov}(X, Y) = \mathsf{E}(XY) - \mathsf{E}(X)\,\mathsf{E}(Y)$$

$$= \frac{pQ}{(q-Q)^2} + \frac{pq}{(q-Q)^2} - \frac{pq}{(P-p)(q-Q)} = \frac{pQ}{(q-Q)^2}$$

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} = \frac{pQ}{(q-Q)^2} \cdot \frac{(q-Q)^2}{\sqrt{pq\,PQ}} = \sqrt{\frac{pQ}{qP}}$$

Exercise 8(c)

1. Using p.g.f. find the probability that the sum of eyes on three fair dice will be 9.
2. A die is thrown repeatedly until either a 2 or a 3 is obtained. Show that the p.g.f. of the number of throws required is $t/(3-2t)$.
3. A bag contains four balls numbered 0, 1, 1 and 2. Suppose n balls are successively drawn, with replacement between drawings. Let X be the sum of n numbers drawn. Find the p.g.f. of X and hence show that

$$P(X=k)=\binom{2n}{k}\left(\tfrac{1}{2}\right)^{2n} \quad k=0,1,2,\ldots,2n.$$

4. A fair die is thrown 6 times. Let X be the total number of points. Verify that the probability distribution of X is :

k :	6	7	8	9	10	
p_k :	λ	6λ	21λ	56λ	126λ	$(\lambda=1/6^6)$

5. For the p.m.f. : $P(X=a)=a_x\theta^x/f(\theta)$, $x=0,1,2,\ldots,\theta>0$ where $a_x\geq 0$ and $f(\theta)=\Sigma a_x\theta^x$, show that the p.g.f. in terms of $f(\theta)$ is $G(t)=f(\theta t)/f(\theta)$.
6. Balls are successively distributed among 3 cells. Let X be the number of balls that must be distributed to occupy all three cells. Show that the p.g.f. of X is $2t^3/(3-2t)(3-t)$.
7. Let X be the number of unoccupied boxes when n balls are randomly distributed among m boxes. Show that

$$G(t:X)=\sum_{j=0}^{m}\binom{m}{j}\left(\frac{m-j}{m}\right)^n (t-1)^j.$$

Hence, show that the p.m.f. of X is given by

$$P(X=k)=\sum_{j=0}^{m-k}(-1)^j\binom{k+j}{j}\binom{m}{k+j}\left(1-\frac{k-j}{m}\right)^n.$$

8. Let $X_1, X_2, \ldots$ be i.i.d. variates, each with distribution : $P(X_1=0)=P(X_i=1)=P(X_i=2)=1/3$. Tabulate the distribution of : S_1, S_2, S_3, S_4 where

$S_n=X_1+X_2+\ldots+X_n$. Verify the following expression : $[\lambda=(1/3)^n]$

$P(S_n=0)=\lambda,\ P(S_n=1)=n\lambda,\ P(S_n=2)=\lambda\binom{n+1}{2}$

9. Let p_n be the probability that in n tosses of an ideal coin, no run of three (consecutive) heads appears. Show that

$$Q(t)=\Sigma p_n t^n=(2t^2+4t+8)/(8-4t-2t^2-t^3).$$

Explain the use of the real zero of the denominator in getting an approximate value of p_n for not too small a value of n. Obtain a value of $Q(t)$ in case of runs of k heads.

10. In a sequence of Bernoulli trials, let P_n be the probability of an even number of successes. Prove the recurrence formula : $P_n=qP_{n-1}+(1-P_{n-1})p$.

From this derive the generating function and hence obtain the explicit formula for P_n.

11. In a sequence of Bernoulli's trials, let P_n be the probability that the first combination SF occurs at trials number $(n-1)$ and n. Find the generating function, mean and variance.

12. A series of independent Bernoulli's trials is performed until an un-interrupted run of r successes is obtained for the first time where r is a given positive integer. Assuming that the probability of a success in any trial is p, $(q = 1 - p)$, show that the p.g.f. of the number of trials is $G(t) = p^r t^r (1 - pt)/(1 - t + qp^r t^{r+1})$.

13. Show that the only way that $G(t)$ and $1/G(t)$ can both be p.g.f. is that $G(t)$ is identically 1.

14. "There is a unique one-to-one relation between the p.m.f. of a non-negative integer-valued random variable and its p.g.f.". Prove this assertion.

15. Let $G(t) = \mathsf{E}(t^X)$. For any integer-valued variate X, show that

$$\sum_{n=0}^{\infty} P(X > n)\, t^n = \frac{1 - G(t)}{1 - t};\quad \sum_{n=0}^{\infty} P(X \le n)\, t^n = \frac{G(t)}{1 - t},\quad \text{Deduce : } \lim_{t \to 1} \frac{1 - G(t)}{1 - t} = \mathsf{E}(X).$$

Show also that, if $0 < \mathsf{E}(X) < \infty$, then $[1 - G(t)]/[(1 - t)\,\mathsf{E}(X)]$ is a p.g.f.

16. Express μ_4' in terms of $G(t)$, and also in terms of $H(t) = \Sigma P(X > n) t^n$.

17. Let $G(t_1, t_2)$ be the p.g.f. of a pair of integer-valued r.v.'s Prove :
    (a) The p.g.f.'s of marginals $P(X = j)$ and $P(Y = k)$ are $G(t, 1)$ and $G(1, t)$ respectively.
    (b) The p.g.f. of $X + Y$ is $G(t, t)$.
    (c) X and Y are independent iff $G(t_1, t_2) = G(1, t_2)\, G(t_1, 1)$.

18. The p.g.f. of X and Y is given by $G(s, t) = \exp[-\lambda - \mu - b + \lambda s + \mu t + bst]$.
    (a) Obtain the marginal p.f.g.'s and identify them.
    (b) Obtain the p.g.f. of $X + Y$ and show that $P(X + Y = 0) = \exp[-(\lambda + \mu + b)]$.
    (c) Show that the case $b = 0 \Rightarrow X$ and Y are indep.

19. Let $X \sim$ Pois (λ) and $Y \sim$ Pois (μ) be independent. Show that the joint p.g.f. of X and $Z = X + Y$ is $G(t_1, t_2) = \exp\{\lambda(t_1 t_2 - 1) + \mu(t_2 - 1)\}$.
    Obtain the marginal distribution of Z and show that Cond. dist. of X, given $Z = n$ is

$$P(X = r \mid Z = n) = \binom{n}{r} \left(\frac{\rho}{1 + \rho}\right)^r \left(\frac{1}{1 + \rho}\right)^{n-r},\quad 0 \le x \le n,\ \left[\rho = \frac{\lambda}{\mu}\right].$$

8-80. *Mellin Transform*

Definition. The Mellin Transform, $T(t)$, of a variate X is defined by

$$T_X(t) = \mathsf{E}(X^t),\ X \ge 0\,.$$

for all values of t for which $\mathsf{E}(X^t)$ exists.

Connection. $T(t : X) = \mathsf{E}(X^t) = \mathsf{E}(e^{t \log x}) = M(t : \log X)$. Thus $T_X(t) = M_{\ln X}(t)$

Remark. Mellin Transform is moment of arbitrary order of a positive r.v.

Theorem. If X and Y are independent variates, then

$$T(t : XY) = T(t : X)\,.\,T(t : Y).$$

Proof. $T(t : XY) = M(t : \ln XY) = M(t : \ln X + \ln Y) = M(t : \ln X)\,.\,M(t : \ln Y) = T(t : X)\,T(t : Y)$.

Mellin Transformers of B_I Distribution and Gamma Distribution

(i) Let $Z \sim B_1(a, b)$, then $f(z) = z^{a-1}(1-z)^{b-1} / B(a, b), 0 < z < 1, a > 0, b > 0$.

$$\therefore \quad T(\theta : Z) = E(Z^\theta) = \int_0^1 z^\theta \frac{z^{a-1}(1-z)^{b-1}}{B(a,b)} dz = \frac{B(a+\theta, b)}{B(a,b)} = \frac{\Gamma(a+\theta)}{\Gamma(a)} \cdot \frac{\Gamma(a+b)}{\Gamma(a+b+\theta)}$$

$$= a^{[\theta]} / (a+b)^{[\theta]}. \quad [a^{[\theta]} = a(a+1)\dots(a+\theta-1)]. \text{ Reverse Factorial}$$

(ii) If $Z \sim$ gam $(\lambda; k)$ then

$$T(\theta : Z) = E(Z^\theta) = \int_0^\infty \frac{z^\theta (\lambda)^k e^{-\lambda z} z^{k-1}}{\Gamma(k)} dz = \frac{\Gamma(k+\theta)}{(\lambda)^\theta \Gamma(k)} = \frac{k^{[\theta]}}{(\lambda)^\theta}.$$

***Example* 1.** Let $X \sim B_1(a, b), Y \sim B_1(a+b, c)$ and $Z \sim$ gam $(\lambda, a+b)$ be independent distributed Find the distributions of XY and XZ.

Solution. (i) We utilize Mellin Transform technique :

$$T(\theta : XY) = T(\theta : X)\, T(\theta : Y) = \frac{a^{[\theta]}}{(a+b)^{[\theta]}} \frac{(a+b)^{[\theta]}}{(a+b+c)^{[\theta]}} = \frac{a^{[\theta]}}{(a+(b+c))^{[\theta]}}$$

By Continuity (Uniqueness) Theorem, we conclude that $XY \sim B_1(a, b+c)$.

(ii) $$T(\theta : XZ) = T(\theta : X)\, T(\theta : Z) = \frac{a^{[\theta]}}{(a+b)^{[\theta]}} \frac{(a+b)^{[\theta]}}{\lambda^{[\theta]}} = \frac{a^{[\theta]}}{\lambda^{[\theta]}}$$

By Continuity Theorem, $XZ \sim$ gam (λ, a).

Note. If $X \sim B_{11}(a, b), T_X(t) = \dfrac{\Gamma(a+t)\,\Gamma(b-t)}{\Gamma(a)\,\Gamma(b)} = \dfrac{a^{[t]}}{(b-t)^{[t]}}$

Exercise. Solve this problem by Transformation of variables.

***Example* 2.** Find the Mellin Transform of log-normal distribution " Show that, if $X \sim$ L-N (μ_1, σ_1^2), $Y \sim$ L-N (μ_2, σ_2^2) are independent then XY ~ L-N $(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$

Solution. If $U \sim N(\mu, \sigma^2)$, then $V = e^U \sim$ L-N (μ, σ^2) Hence

$$T(t : V) = E(V^t) = E(e^{tU}) = e^{\mu t + (\sigma^2 t^2/2)}, \qquad [\text{m.g.f. of } N(\mu, \sigma^2)]$$

$$T(t : XY) = T(t : X)\, T(t : Y) = \exp[(\mu_1 t + (\sigma_1^2 t^2/2)] \exp[\mu_2 t + (\sigma_2^2 t^2/2)]$$

$$= [(\mu_1 + \mu_2)t + \tfrac{1}{2}(\sigma_1^2 + \sigma_2^2)t^2]$$

Thus $XY \sim$ L-N $(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$ as gauranteed by Continuity Theorem

***Life is sexually transmitted disease.* (R.D. Laung)**

The most difficult thing is to open a closed mind.

Characteristic Functions 9

9-10. *Definition : Characteristics Function*

The characteristic function of a random variable X is denoted by $\phi(t : X)$ or $\phi_X(t)$, equivalently $M_X(it)$, and is defined by the relation

$$M_X(it) \equiv \phi(t : X) = \mathsf{E}(e^{itX}). \qquad \left(i = \sqrt{-1}\right). \qquad ...(1)$$

Existence. The m.g.f. does not always exist, and sometimes it exists only for small values of t. However, $\phi_X(t)$ always exists since

$$| M(it) | = | \mathsf{E}(e^{itX}) | \le \mathsf{E}\,(| \cos tX + i \sin tX |) = 1.$$

Connection. $\qquad M(t) = G(e^t) = \phi(-it)\ ; \quad \phi(t) = M(it) = G(e^{it}).$

Note. The word 'characteristic function' shall sometimes be written as 'Ch. function', and $\phi(t : X)$ will often be written simply as $\phi(t) = M(it)$.

Effect of Linear Transformation on $\phi(t)$

$$M(it : a + bX) = e^{ait}\, M(ibt : X)$$

Proof. $M(it : a + bX) = \mathsf{E}[e^{it(a+bX)}] = \mathsf{E}[e^{(ia)t} . e^{i(bt)X}] = e^{iai}\, \mathsf{E}[e^{i(bt)X}]$ [by Def. & by Lin E]

$$= e^{iat}\, M(ibt : X). \qquad \text{[by Def. of Ch. Function]}$$

Note. $\qquad M(it : X - \mu) = e^{-\mu it}\, M\,(it : X).$

Linear Combination Theorem

If X and Y are *independent* variates, and a, b, c any constants,then

$$M(it : aX + bY + c) = e^{itc}\, M(ait : X)\, M(bit : Y).$$

Proof. $M(it : aX + bY + c) = \mathsf{E}\{e^{it(aX+bY+c)}\} = \mathsf{E}\{e^{itaX}\, e^{itY}\, e^{itc}\}$ (Definition)

$$= e^{itc}\, \mathsf{E}(e^{itaX})\, \mathsf{E}(e^{ibtY}) = e^{itc}\, M(iat : X)\, M\,(ibt : Y) \quad \text{[by Ind}(X, Y)\text{ and Lin E]}$$

In particular : $\quad M(it : X + Y) = M(it : X)\, M(it : Y)$. [Addition Theorem]

9-11. *Moments through Ch. Function*

If $\mathsf{E}(X^r) = \mu'_r$, $1 \le r \le n$, exist, then $\phi_X(t)$ is n times differentiable and $\phi^{(r)}(0) = (i)^r\, \mu'_r$, $1 \le r \le n$.

Proof. Let $F(x)$ be the c.d.f. of X. Since μ'_r exists, so

$$\int_{-\infty}^{\infty} | x |^r\, dF(x) < \infty \Rightarrow \int_{-\infty}^{\infty} x^r\, e^{itx}\, dF(x) < \infty \qquad ...(1)$$

converges uniformly in t.

Hence, differentiating the Ch. Function r times [which is permissible by (1)] we get

$$\phi^{(r)}(t) = (i)^r \int_{-\infty}^{\infty} x^r e^{itx}\, dF(x) \Rightarrow \phi^{(r)}(0) = (i)^r \mu_r', \text{ i.e. } \mu_r' = (-i)^r \phi^{(r)}(0), \quad 1 \le r \le n$$

Cor. Maclaurin's expansion of the function $\phi_X(t)$ is

$$\phi(t) = \sum_{r=0}^{n} \mu_r' \frac{(it)^r}{r!} + O(t)^n .$$

As $t \to 0$, μ'_r = coefficient of $(it)^r/r\,!$

Note. **Crucial comments on evaluation of complex integrals**

To avoid using calculus of residues in calculating Ch. Functions, by writing it for t in m.g.f., is not rigorous unless justified. It produces correct result for normal and gamma distributions, as well as many others, but may not succeed with several others, e.g. Cauchy distribution (which possesses no m.g.f.). Do not fall into the trap of treating i as if i is a real number, even though this malpractice yields the correct answer in some cases. However, we accept the formal procedure :

If $M_X(t)$ is finite in a non-trivial nbd of origin, we accept $\phi_X(t) = M_X(it)$.

9-12. *Worked-out Problems*

***Example* 1.** Show that if X and Y are independent, then $\phi(t : X + Y) = \phi(t : X) \cdot \phi(t : Y)$. But the converse need not be true.

Solution. $\phi(t : X + Y) = E[e^{it(X+Y)}] = E[e^{itX} \cdot e^{itY}] = E(e^{itX}) \cdot E(e^{itY})$ [$\because$ X & Y are indep.]

This result is a special case of linear combination theorem.

Converse. (i) Consider the Cauchy random variable X having p.d.f.

$$f(x) = 1/\pi\,(1 + x^2), \quad -\infty < x < \infty.$$

This is known to have the Ch. Function $\phi(t : X) = e^{-|t|}$ [§9-70]

Let $X_1 = aX$, $X_2 = bX$, $a > 0$, $b > 0$, then $X_1 + X_2 = (a + b)X$.

$$\phi(t : X_1) = e^{-a|t|}, \phi(t : X_2) = e^{-b|t|}, \phi(t : X_1 + X_2) = e^{-(a+b)|t|} \Rightarrow \phi(t : X_1 + X_2) = \phi(t : X_1)\, \phi(t : X_2)$$

However, X_1 and X_2 are not independent; they are correlated : $bX_1 = aX_2$.

(ii) For another example, see Worked-out Problem 3.

***Example* 2.** Find m.g.f. and Ch. Function of the triangular density

$$f(x) = 1 - |x|, \quad |x| \le 1, \quad f(x) = 0, \text{ otherwise.}$$

Solution. Here $f(x) = (1 + x)\, I(-1 \le x \le 0) + (1 - x)\, I(0 \le x \le 1) + 0\, I(|x| > 1)$.

$$\therefore M_X(t) = E(e^{tX}) = \int_{-1}^{0} (1 + x) e^{tx}\, dx + \int_0^1 (1 - x) e^{tx}\, dx = \int_0^1 (1 - z) e^{-tz}\, dz + \int_0^1 (1 - x) e^{tx}\, dx \; [z = -x]$$

$$= \int_0^1 (1 - x)(e^{-tx} + e^{tx})\, dx = 2\int_0^1 (1 - x) \cosh tx\, dx. \qquad \text{[Integ. by Parts]}$$

$$= 2\left[(1 - x)\left(\frac{\sinh tx}{t}\right) - (-1)\left(\frac{\cosh tx}{t^2}\right)\right]_0^1 = 2\frac{(\cosh t - 1)}{t^2}$$

So $\phi(t) = M(it) = 2(1 - \cos t)/t^2$.

Since $$M(t) = \frac{2}{t^2}\sum_{r=1}^{\infty}\frac{t^{2r}}{(2r)!} = 2\sum_{r=1}^{\infty}\frac{t^{2r-2}}{(2r)!}$$

it follows that $E(X) = 0$, [= coeff. of t] ; Var $(X) = 1/6$ [= coeff. of $t^2/2$!].

***Example* 3.** Let the joint-density of random vector (X, Y) be :

$f(x, y) = (4a^2)^{-1}\,[1 + xy(x^2 - y^2)]$, $|x| \le a$, $|y| \le a$, $a > 0$. $f(x, 0) = 0$, elsewhere.

Show that $\phi(t : X + Y) = \phi(t : X)\,\phi(t : Y)$, although X and Y are dependent variates.

Solution. The marginal densities are trivially determined, as odd-part of integrand yields zero.

$$f_X(x) = \int_{-a}^{a}\frac{1}{4a^2}[1 + xy(x^2 - y^2)]\,dy = \frac{1}{2a} \cdot f_Y(y) = \int_{-a}^{a}\frac{1}{4a^2}[1 + xy(x^2 - y^2)]\,dx = \frac{1}{2a}$$

Since $f_X(x) \cdot f_Y(y) \ne f(x, y)$, it follows that X and Y are not independent.

$$\phi(t : X) = \int_{-a}^{a}\frac{e^{itx}}{2a}\,dx = \int_{-a}^{a}\left(\frac{\cos tx + i\sin tx}{2a}\right) = \frac{1}{a}\int_0^a \cos tx\,dx = \left(\frac{\sin at}{at}\right)$$

By symmetry, $\phi(t : Y) = \phi(t : X)$ $\sin at/(at)$.

We now find the density of $U = X + Y$. Let $x + y = u$, $x - y = v$ so that

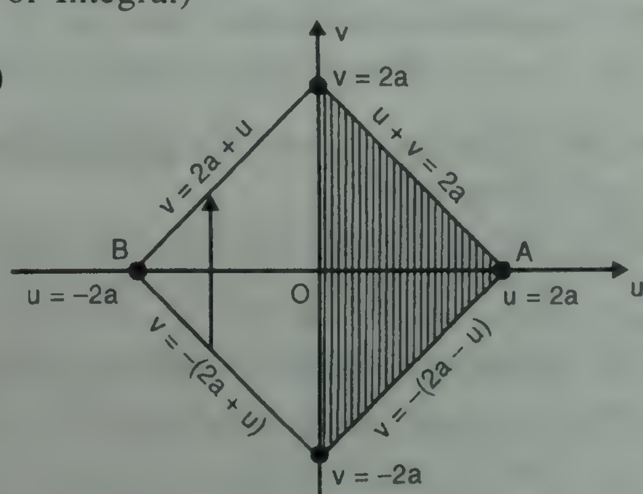
$$x = \frac{1}{2}(u + v),\ y = \frac{1}{2}(u - v),\ \left|\frac{\partial(x, y)}{\partial(u, v)}\right| = \frac{1}{2},\quad \text{i.e.}\quad dx\,dy = \frac{1}{2}\,du\,dv.$$

The joint p.d.f. of u, v is, thus, $g(u, v) = (8a^2)^{-1}\left[1 + \frac{1}{4}uv\,(u^2 - v^2)\right]$.

Note that $x = \pm a \;\Rightarrow\; u + v \pm 2a$, $y = \pm a \;\Rightarrow\; v - u = \pm 2a$.

The density of u is, (use Even-Odd properties of Integral)

$$g_1(u) = \begin{cases}\int_{-(2a-u)}^{2a+u} g(u, v)\,dv \\ \int_{-(2a-u)}^{2a+u} g(u, v)\,dv\end{cases} = \begin{cases}\dfrac{u + 2a}{4a^2}, & -2a \le u \le 0 \\[2mm] \dfrac{2a - u}{4a^2}, & 0 \le u \le 2a\end{cases}$$



$$\therefore\ \phi(t : u) = \int_{-2a}^{0} e^{itu}\left(\frac{u+2a}{4a^2}\right)du + \int_0^{2a} e^{itu}\left(\frac{2a-u}{4a^2}\right)du$$

$$= \int_0^{2a} e^{-itz}\left(\frac{2a - z}{4a^2}\right)dz + \int_0^{2a} e^{itu}\left(\frac{2a - u}{4a^2}\right)du$$

$$= \int_0^{2a} e^{itu}\left(\frac{e^{itu} + e^{-itu}}{2}\right)\frac{2a - u}{2a^2}\,du, [z = -u]$$

$$= \int_0^{2a}\left(\frac{2a - u}{2a^2}\right)\cos tu\,du = \left(\frac{1 - \cos 2at}{2a^2 t^2}\right) = \left(\frac{\sin at}{at}\right)^2. \qquad \text{[Integ. by parts]}$$

It follows that, $\phi(t : X + Y) = \phi(t : X)\,\phi(t : Y)$; although X and Y are not independent.

Problems with Solutions Provided at the End of the Text

1*. If $\mu'_r = \Gamma(n+r)/\Gamma(n)$, find $\phi(t)$.

2*. For a distribution, the cumulants are given by

$$k_r = [(r-1)\,!]n, \;\; n > 0.$$

Find the characteristic function.

3*. For the median law : $f(x) = y_0\, e^{-|x|}, -\infty < x < \infty$ find μ, σ M.D. and $\phi(t)$

9-20. *Characteristic Functions of Some Key Distributions*

Evaluations of Ch. Functions are similar to those of m.g.f. However, the correct evaluations in the case of continuous variates needs calculus of residues. We record results, using e^{it} instead of e^{t} in m.g.f.; proofs are then identical with those of m.g.f.

1. ***Binomial distribution*** : bin (n, p) . $\phi(t) = (q + pe^{it})^n$.
2. ***Poisson distribution*** : Pois (λ) . $\phi(t) = e^{\lambda(e^{it}-1)}$.
3. ***Geometric distribution*** : geom (p); $\phi(t) = p/(1 - qe^{it})$; gem(p) ; $\phi(t) = pe^{it}/(1 - qe^{it})$
4. ***Neg-bin distribution*** : $NB(k, p)$; $\phi(t) = [p/(1 - qe^{it})]k$,
   Pascal N-B distribution : $NB^*\,(k, p)$, $\phi(t) = [pe^{it}/(1 - qe^{it})]^k$.
5. ***Multinomial distribution*** : $M(n; k, p_i)$, $\phi(t) = (p_1 e^{it_1} + p_2 e^{it_2} + \ldots + p_k e^{it_k})^n$.
6. ***Uniform distribution*** : unif(a, b) . $\phi(t) = (e^{ibt} - e^{iat})/it\,(b - a)$.
7. ***Normal (Gaussian distribution)*** : $N(\mu, \sigma^2)$, $\phi(t) = e^{i\mu t - \sigma^2 t^2/2}$; $N(0,1): \phi(t) = e^{-t^2/2}$.
8. ***Gamma distribution*** : gam (a, λ), $\phi(t) = (1 - it/\lambda)^{-a}$,
   Expo (λ) : $\phi(t) = (1 - it/\lambda)^{-1}$; Chi-square : $\chi^2_{(n)}$ $\phi(t) = (1 - 2it)^{-n/2}$.
9. ***Laplace distribution*** : Lap (μ, λ), $\phi(t) = e^{i\mu t}\,(1 + t^2/\lambda^2)^{-1}$.
10. ***Cauchy distribution*** : Chy (a, b), $\phi(t) = e^{i\mu t - b|t|}$. [§9-70]

9-30. *Necessary but not Sufficient Conditions for a Function to be a Ch. Function*

If $\phi(t) \equiv M(it)$ is the Ch. Function of the variate X, then

(i) $\phi(0) = 1$, $M(-it) = \overline{M(it)}$, $|\,\phi\,(t)\,| \le 1$.

(ii) $M(it)$ is uniformly continuous, for all values of t.

Proof. By definition, $\phi(t) = \mathrm{E}(e^{itX})$.

(i) $\phi(0) = \mathrm{E}(e^{ioX}) = \mathrm{E}(1) = 1$.

$\phi(-t) = \mathrm{E}(e^{-itX}) = \mathrm{E}\,(\cos tX - i \sin tX) = \mathrm{E}\,(\cos tX) - i\mathrm{E}\,(\sin tX)$

$\overline{\phi(t)} = \text{conj}\; \mathrm{E}(e^{itX}) = \text{conj}\;[\mathrm{E}\,(\cos tX) + i\mathrm{E}\,(\sin tX)] = \mathrm{E}\,(\cos tX) - i\mathrm{E}\,(\sin tX)$

$\therefore$ $\phi(-t) = \overline{\phi(t)}$. [m.g.f. does not possess this property]

Also $|\,\phi(t)\,| = |\,\mathrm{E}(e^{itX})\,| \le \mathrm{E}\,(|\,e^{itX}\,|) = 1$.

(ii) ***Uniform Continuity.***

For $\delta > 0$, define a number b such that

$$P(X > b) = 1 - F_X(b),\ P\{X \le -b\} = F_X(-b) < \varepsilon/4$$

Construct the difference Δ as under

$$\Delta = |M_X[i(t+\delta)] - M_X(it)| = \left|\int_{-\infty}^{\infty} (e^{i(t+\delta)x} - e^{itx})\, dF(x)\right|$$

$$= \left|\int_{-\infty}^{\infty} e^{itx}(e^{i\delta x} - 1)\, dF(x)\right| \le \int_{-\infty}^{\infty} |e^{i\delta x} - 1|\, dF(x)$$

As $|1 - e^{-i\delta x}| = 2\left|\sin\left(\frac{\delta x}{2}\right)\right| \le 2$, the above yields

$$\Delta \le 2\left\{\int_{-\infty}^{-b} dF(x) + \int_{-b}^{b} \left|\sin\frac{\delta x}{2}\right| + \int_{b}^{\infty} dF(x)\right\}$$

The mid-integral is independent of t that can be made arbitrarily small by making δ arbitrarily small. Hence

$\lim\limits_{\delta \to 0} \Delta \le \varepsilon$, uniformly in t and so $M_X(it)$ is uniformly continuous in $]-\infty, \infty[$

Illustrations : (i) $\psi(t) = \left(\frac{3}{10} e^{it} + 1\right) e^{5(e^{it}-1)}\, e^{5it - t^2}$ (ii) $\psi(t) = e^{4it + t^2}$

(i) This is not a characteristic function because $\psi(0) \ne 1$

(ii) This is not a Ch. Function because $|\psi(t)| = e^{t^2} \nless 1$.

9-31. *Symmetry Property*

If a variate X is symmetric about 0, i.e. if $f(-x) = f(x)$, then $M(it)$ is real and even function of t. Conversely, if $M(it)$ is real, then X is symmetric about zero.

Proof. $M(it) = \mathrm{E}[e^{itX}] = \mathrm{E}[(\cos tX) + i \sin(tX)] = \mathrm{E}(\cos tX) + i\mathrm{E}(\sin tX)$, by L in E

Since, $\quad \mathrm{E}(\sin tX) = \int_{-\infty}^{\infty} f(x) \sin tx\, dx = 0,\quad$ [Odd Integrand ; $f(-x) = f(x)$]

$\therefore \quad M(it) = \mathrm{E}(\cos tX).$

Thus, whenever X is symmetric, $\phi(t)$ is real and even function of t.

Converse. Let $\phi_0(t)$ be a real Ch. Function. Set $Y = -X$, then

$$\phi_0(t : Y) = \phi_0(t : -X) = \phi_0(-t : X) = \overline{\phi_0(t : X)} = \phi_0(t : X)$$

because ϕ_0 is real. Thus $\phi(t : Y) = \phi_0(t : X)$, so that their distribution functions must be the same

Now $\quad F_Y(y) = P(Y \le y) = P(X \ge -y) = 1 - P(X \le -y) \Rightarrow F_Y(y) = 1 - F_X(-y)$

Since $\quad F_Y(y) = F_X(y)$ so $F_X(y) = 1 - F_X(-y)$,

which on differentiation gives $f_X(y) = f_X(-y)$, whence X is symmetric about $X = 0$

9-32. *Some Interesting Inequalities*

Let $M_X(it) = M_1(t) + iM_2(t)$. Then

1. $1 - M_1(2t) \le 4\,[1 - M_1(t)]$.
2. $1 - |\,M(2it)\,|^2 \le 4\,[1 - |\,M(it)\,|^2]$.
3. $|\,M(it) - M(it + ih)\,|^2 \le 2\,|1 - M_1(h)\,|$, t and h real.
4. $|\,M_2(t)\,| \le [1 - M_1(2t)]^{1/2}$.

Proof. **1.**

$$M_1(t) = \int_{-\infty}^{\infty} \cos tx\, f_X(x)\,dx \qquad \text{[Def.]} \qquad \ldots\text{(i)}$$

$$1 - \cos 2tx = 2\sin^2 tx = 2\,(1 - \cos tx)\,(1 + \cos tx) \le 4\,(1 - \cos tx).$$

Multiply this inequality by $f(x)$ and integrate over $]-\infty, \infty[$; then use (i) to recover

$$1 - M_1(2t) \le 4\,[1 - M_1(t)].$$

2. This result follows on replacing $[M_1(2t)]^2$ by $|\,M(2it)\,|^2$ in (1).
The reason is that $|\,M(it)\,|^2$ is a real-valued Ch. Function.

3.

$$|\,M(it) - M(it + ih)\,|^2 = \left|\int e^{itX}(1 - e^{iXh})\,f(x)\,dx\right|^2, \quad -\infty < x < \infty$$

$$\le \left\{\int_{-\infty}^{\infty} \left|e^{itx}\sqrt{f(x)}\right|^2 dx\right\}.\left\{\int_{-\infty}^{\infty} \left|(1 - e^{ixh})\sqrt{f(x)}\right|^2 dx\right\}, \quad \text{[by C-S inequality §10-14]},$$

Since $\int_{-\infty}^{\infty} \left|e^{itx}\right|^2 \left|\sqrt{f(x)}\right|^2 dx = \int_{-\infty}^{\infty} f(x)\,dx = 1, \quad |(1 - e^{ixh})| = 2(1 - \cos xh);$

$$\therefore \qquad M(it) - M(it + ih) \le \int_{-\infty}^{\infty} 2(1 - \cos xh)\,f(x)\,dx = 2\,[1 - M_1(h)].$$

If X is discrete, the proof is similar.

Note. This result also proves that $M(it)$ is uniformly continuous over the entire range of t.

4.

$$|\,M_2(t)\,| = \left|\int_{-\infty}^{\infty} \sin tx\, dF(x)\right| \le \left[\int_{-\infty}^{\infty} \sin^2 tx\, dF(x)\right]^{1/2}$$

$$= \left(\frac{1}{\sqrt{2}}\right)\left[\int_{-\infty}^{\infty} (1 - \cos 2tx)\,dF(x)\right]^{1/2} = \left(\frac{1}{\sqrt{2}}\right)[1 - M_1(2t)]^{1/2}$$

$$\le [1 - M_1(2t)]^{1/2}$$

9-33. *Bochner's Test*

A function $\phi(t)$ of a real variable t on the interval $]-\infty, \infty[$, is a Ch. Function of a Dist. iff

(i) $\phi(0) = 1$, (ii) $\phi(t)$ is continuous

(ii) $\phi(t)$ is non-negative definite : $\sum_{j=1}^{\infty} \sum_{k=1}^{\infty} \phi(t_j - t_k)\, w_j\, \overline{w}_k \ge 0$

for any n, any set $t_1, t_2, \ldots, t_n$ of real numbers, and any set $w_1, w_2, \ldots, w_n$ of complex numbers.

Proof. **Necessity :** If $\phi(t)$ is a Ch. Function, then necessary conditions (i) and (ii) have already been proved [§9-30]. We now attend to quadratic form in (iii). The definition $\phi(t) = \mathsf{E}(^{itX})$ provides

$$\phi(t_j - t_k) = \mathsf{E}[e^{i(t_j - t_k)X}].$$

$$\therefore \quad 0 \le \mathsf{E}[\,|\Sigma w_k e^{it_k X}|^2\,] = \mathsf{E}\left[\left(\sum_{j=1}^{n} w_j\, e^{it_j X}\right)\left(\sum_{k=1}^{n} \overline{w}_k\, e^{-it_k X}\right)\right] = \mathsf{E}\left\{\sum_{j=1}^{n}\sum_{k=1}^{n} w_j\, \overline{w}_k\, e^{i(t_j - t_k)X}\right\}$$

$$= \sum_{j=1}^{n}\sum_{k=1}^{n} w_j\, \overline{w}_k\, \mathsf{E}\left[e^{i(t_j - t_k)X}\right] = \sum_{j=1}^{n}\sum_{k=1}^{n} w_j\, \overline{w}_k\, \phi(t_j - t_k).$$

Sufficiency. We omit the proof of this part.

9-40. *Periodic Aspect of Integer-valued Variates*

We show that $P(X \in \mathbb{Z}) = 1, \;\Rightarrow\; M_X(it) = M_X[i(t + 2\pi)]$ [Period $= 2\pi$]

Converse holds at $t = 0$, i.e. $M_X(2\pi i) = 1 \;\Rightarrow\; P\{X \in \mathbb{Z}\} = 1$.

(*i*) Here $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$. For such elements we need to show that $M(it)$ is periodic with period 2π. Writing $p_k = P(X = k)$, we have

$$\sum_{k=-\infty}^{\infty} p_k = \sum_{-\infty}^{\infty} P(X = k) = P\{X \in \mathbb{Z}\} = 1.$$

$$M_X\,[i(t + 2\pi)] = \mathsf{E}\left[e^{i(t+2\pi)X}\right] = \sum_{k=-\infty}^{\infty} p_k\, e^{i(t+2\pi)k} \quad \left[(e^{2\pi i})^k = 1\right]$$

$$= \sum_{-\infty}^{\infty} p_k\, e^{itk} = \mathsf{E}\left(e^{itX}\right) = M_X(it).$$

This shows that $M(it)$ is periodic with period 2π.

(ii) $M_X(2\pi i) = \mathsf{E}\left(e^{2\pi i X}\right) = \mathsf{E}\left\{\cos 2\pi X + i \sin 2\pi X\right\} = 1,$ (by hypothesis)

Equating real parts : $\mathsf{E}\left(\cos 2\pi X\right) = 1 \;\Rightarrow\; \mathsf{E}\left\{1 - \cos 2\pi X\right\} = 0 \;\Rightarrow$

$\mathsf{E}\left(2 \sin^2 \pi X\right) = 0$. This holds if $\sin \pi X = 0 = \sin k\pi$. Thus $X = k$, where $k = 0, \pm 1, \pm 2$. This is what $P\{X \in \mathbb{Z}\} = 1$ stands for.

Integer-valued Inversion Formula

If $P\{X \in \mathbb{Z}\} = 1$, we show that

$$P\{X = n\} = \frac{1}{2\pi}\int_0^{2\pi} e^{-int}\, M_X(it)\, dt. \qquad \ldots(1)$$

Illustrate the distribution of X when $M_X(it) = \cos\,(t/2)$.

Proof. Since $M(it) = \Sigma e^{itk} p_k, n \in \mathbb{Z}$, a typical term in R.H.S. of (1) is

$$\int_0^{2\pi} e^{-int}\, e^{itk}\, p_k\, dt = \int_0^{2\pi} e^{it(k-n)} p_k\, dt = \begin{cases} 0, & \text{if } k \ne n \\ 2\pi\, p_k, & \text{if } k = n \end{cases} = 2\pi\, p_k\, \delta_{kn}$$

This is trivial : $\int_0^{2\pi} (\sin\theta, \cos\theta)\, d\theta = (0, 0)$

Assume that summation and integration can be interchanged*, then

$$\int_0^{2\pi} e^{-int} M_X(it)\,dt = \int_0^{2\pi} e^{-int}\left(\sum_{k\in\mathbb{Z}} e^{itk} p_k\right)dt = \sum_{k\in\mathbb{Z}} \int_0^{2\pi} e^{it(k-n)} p_k\,dt$$

$$= \sum_{k\in\mathbb{Z}} 2\pi\, p_k\, \delta_{nk} = 2\pi\, p_n$$

This yields the formula (1).

Second Part : $M_X[i\,(t + 4\pi)] = M_X(it)$, it follows that $M_X\,(it)$ is periodic with period 4π. For period 2π, let $Y = 2X$. Then

$M(it : Y) = M(2it : X) = \cos t$ [periodic with period 2π]. It follows that $P\{Y \in \mathbb{Z}\} = 1$. To determine distribution of X, i.e. $P(X = n) = p_n$, we go via a distribution of Y.

Now $$P\left(X = \frac{1}{2}\right) = P(Y = 1) = \frac{1}{2\pi}\int_0^{2\pi} e^{-it} M_Y(it)\,dt \qquad \text{[by (1) with } n = 1]$$

$$= \frac{1}{2\pi}\int_0^{2\pi} e^{-it} \cos t\,dt \quad \left[e^{-it} = \cos t - i \sin t\right]$$

$$= \frac{1}{2\pi}\int_0^{2\pi} \cos^2 t\,dt = \frac{1}{2}$$

Similarly, $P(X = -1/2) = P(Y = -1) = 1/2$.

Thus $P\{X = \pm\, 1/2\} = 1/2$, so $P\{X = n/2\} = P\{Y = n\} = 0$ if $n \notin \{-1, 1\}$.

9-41. *Ch. Function of Mixtures and Compounds of Distributions*

Given that $\phi, \phi_1, \phi_2, \phi_3, \ldots$ are Ch. Function. Then,

1. *Finite linear combination* : $a\phi_1 + b\phi_2$ is a Ch. Function, if $a \ge 0$, $b \ge 0$, $a + b = 1$.
2. $\phi(-t), \overline{\phi}(t)$ and $\operatorname{Re}\phi(t) = \phi_0(t)$, are all Ch. Functions.
3. Product of Ch. Functions : $\phi_1 \cdot \phi_2 \ldots \phi_k$ is a Ch. Function ; $|\,\phi(t)\,|^2$ is a Ch. Function.
4. The polynomial $a_1\phi + a_2\phi^2 + \ldots + a_n\phi_n$ is a Ch. Function provided $a_i \ge 0$ and $\sum a_i = 1$.
5. Infinite linear combination : $\sum a_i\, \phi_i$, $1 < i \le \infty$ is Ch. Function provided $a_i \ge 0$ and $\sum a_i = 1$.
6. If $G(t)$ is a p.g.f. then $G(\phi(t))$ is a Ch. Function.

---

*Let $A = \int_0^{2\pi} (\Sigma_B\, e^{i(k-n)t} p_k)\,dt, \quad B = \{k \in \mathbb{Z}\}$

$$A_N = \sum_{k=-N}^{N} \int_0^{2\pi} e^{i(k-n)t}\, p_k\,dt = \int_0^{2\pi}\left\{\sum_{k=-N}^{N} e^{i(k-n)t}\, p_k\right\}dt \qquad \text{[Finite sums commute integrals]}$$

We want to show that $A_N \to A$ as $N \to \infty$. So consider

$$|\,A_N - A\,| = \left|\int_0^{2\pi} \sum_{|k|>N} e^{i(k-n)t}\, p_k\,dt\right| \le \int_0^{2\pi} \sum_{|k|>N} \left|e^{i(k-n)t}\, p_k\right| dt = 2\pi\left(\sum_{|k|>N} p_k\right) \to 0, \text{ as } N \to \infty.$$

The last result is obvious, as it is tail of a convergent series $\sum p_k = 1$, $k \in \mathbb{Z}$.

Proof. **1.** Let $\phi(t) = a\phi_1(t) + b\phi(t)$, then $\phi(0) = a\phi_1(0) + bf_2(0) = a + b = 1$.
Since ϕ_1 and ϕ_2 are uniformly continuous, their linear combination is also uniformly continuous.
Since ϕ_1 and ϕ_2 are Ch. Functions. Necessity of Bochner's test gives

$$\sum_{j,k} \phi_r(t_j - t_k)\, w_j\, \overline{w}_k \geq 0, \quad r = 1, 2.$$

$$\therefore \quad \sum_{j,k} \phi_r(t_j - t_k)\, w_j\, \overline{w}_k = a\left\{\sum_{j,k} \phi_1(t_j - t_k)\, w_j\, \overline{w}_k\right\} + b\left\{\sum_{j,k} \phi_2(t_j - t_k)\, w_j\, \overline{w}_k\right\} \geq 0.$$

It follows from Sufficiency of Bochner's test that, $\phi = a\phi_1 + b\phi_2$ is a Ch. Function.
Remark. The proof trivially follows from §9-42.

2. To show that $\overline{\phi}(t)$ is a Ch. Function, we see that the double sum

$$\sum_{j,k} \overline{\phi}(t_j - t_k)\, w_j\, \overline{w}_k = \text{conj}\left[\sum_{j,k} \phi\,(t_j - t_k)\, \overline{w}_j\, w_k\right] \geq 0$$

because $\phi(t)$ is a Ch. Function. Thus $\overline{\phi}(t)$ is also a Ch. Function.

Now, $$\phi_0 = \tfrac{1}{2}\phi + \tfrac{1}{2}\overline{\phi}.$$

Since ϕ and $\overline{\phi}$ are Ch. Functions; their proper linear combination is a Ch. Function.

Note. $\phi_X(-t) = \phi_{(-X)}(t)$; so $\phi(-t)$ is also a Ch. Function. Thus $\phi_0(t) = \frac{1}{2}\phi(t) + \frac{1}{2}\phi(-t)$, being a linear combination of Ch. Functions' is a Ch. Function.
The proofs of the rest of properties all follow, as above, from a direct application of Bochner's Test.
Remarks. Let $\phi_X(t)$ be a Ch. Function. Then $[\phi_X(t)]^{-1}$ is a Ch. Function iff X is degenerate.

9-42. *Convex Combinations of Ch. Functions*

Let $F_1(x)$ and $F_2(x)$ be distribution functions of X_1 and X_2. Let $F_Y(x)$ be defined by

$$F_Y(x) = aF_1(x) + bF_2(x), \quad 0 \leq a, \quad b \leq 1\ ; \quad a + b = 1. \quad \text{[convex combination]}$$

Then, $\phi(t : Y) = a\phi(t : X_1) + b\phi(t : X_2) = a\phi_1(t) + b\phi_2(t)$ [convex combination]

Proof. $\phi(t : Y) = \mathsf{E}(e^{itY}) = \int_{-\infty}^{\infty} e^{ity}\, dF_Y(y) = \int_{-\infty}^{\infty} e^{itx}\, dF_Y(x)$ [dummy y changed]

$$= \int_{-\infty}^{\infty} e^{itx}\,[adF_1(x) + bF_2(x)] = a\int_{-\infty}^{\infty} e^{itx}\, dF_1(x) + b\int_{-\infty}^{\infty} e^{itx}\, dF_2(x) = a\phi_1(t) + b\phi_2(t).$$

Example : Let $X_1 \sim N(0, 1/n)$, $X_2 \sim N(1, 1/n)$ be independent. If variate $Y = X_1$ with probability $1 - (1/n)$, and $Y = X_2$ with probability $(1/n)$; find $\phi(t : Y)$.
Solution. Let $A_j = \{Y = X_j\}$ $j = 1, 2$; then by Multi-Stage Rule

$$P(Y \leq x) = \sum_{j=1}^{2} P(Y \leq x \mid A_j)\, P(A_j) = (1 - n^{-1})\, P(X_1 \leq x) + n^{-1} P(X_2 \leq x). \quad \text{...(1)}$$

Let $Z \sim N(0, 1)$. Then $P(X_1 \leq x) = P(Z \leq x\sqrt{n})$; $P(X_2 \leq x) = P[Z \leq \sqrt{n}(x-1)]$

$$\therefore \quad P(Y \leq x) = \left(1 - \frac{1}{n}\right) F_Z\,(\sqrt{n}x) + \frac{1}{n}\, F_Z\left[\sqrt{n}\,(x-1)\right]$$

So, $\phi_Y(t) = \left(1-\frac{1}{n}\right)\phi_1(t) + \frac{1}{n}\phi_2(t)$ [§9-42. with obvious meanings.]

$$\phi_1(t) = \int_{-\infty}^{\infty} e^{itx}\, dF_Z(x\sqrt{n}) = \int_{-\infty}^{\infty} e^{itz/\sqrt{n}}\, dF(z) = e^{-\frac{1}{2}t^2/n},\ (z = x\sqrt{n})$$

$$\phi_2(t) = \int_{-\infty}^{\infty} e^{itx}\, dF_Z[\sqrt{n}(x-1)] = e^{it}\int_{-\infty}^{\infty} e^{itz/\sqrt{n}}\, dF(z) = e^{it}.e^{-\frac{1}{2}t^2/n},\ [z = \sqrt{n}(x-1)]$$

$$\therefore \quad \phi_Y(t) = e^{-t^2/2n}\left[1-\frac{1}{n}+\frac{1}{n}e^{it}\right].$$

9-43. *Convolution Theorem (Statistical Sum)*

If $\phi_1(t)$ and $\phi_2(t)$ are the characteristics functions of densities $f_1(x)$ and $f_2(x)$, then $\phi(t) = \phi_1(t)\,\phi_2(t)$ is the Ch. Function of their convolution $f(x)$. $[f = f_1 * f_2]$

Proof. By definition of Ch. Function and that of convolution

$$\int_{-\infty}^{\infty} e^{itx} f(x)\, dx = \int_{-\infty}^{\infty} e^{itx}\int_{-\infty}^{\infty} f_1(y) f_2(x-y)\, dy\, dx$$

$$= \int_{-\infty}^{\infty} f_1(y)\int_{-\infty}^{\infty} e^{itx} f_2(x-y)\, dx\, dy \quad \text{[Put } x - y = z \text{ in 2nd part]}$$

$$= \int_{-\infty}^{\infty} f_1(y)\int_{-\infty}^{\infty} e^{it(y+z)} f_2(z)\, dy\, dz = \int_{-\infty}^{\infty} f_1(y)\, e^{ity}\, dy\int_{-\infty}^{\infty} e^{itz}.f_2(z)\, dz = \phi_1(t)\,\phi_2(t).$$

Example : If $\phi_1(t, \theta)$ is a characteristic function, then so is $\int \phi_2(t,\theta)\, dF_1(\theta)$.

Deduce that $\psi(t)$ is also a characteristic function, where $\psi(t) = \frac{1}{t}\int_0^t \phi_1(u)\, du$.

Solution. Let $\phi(t) = \int_{-\infty}^{\infty} \phi_2(t,\theta)\, dF_2(\theta)$.

Obviously, $\phi(0) = 1$ and $\phi(t)$ is a continuous function of t. Let us consider

$$\sum_{j,k}\phi(t_j - t_k)\, w_j\, \bar{w}_k = \sum_{j,k}\left[\int_{-\infty}^{\infty}\phi_2(t_j - t_k,\theta)\, dF_1(\theta)\right] w_j\, \bar{w}_k = \int_{-\infty}^{\infty}\left[\sum_{j,k}\phi_2(t_j - t_k,\theta)\, w_j\, \bar{w}_k\right] dF_1(\theta) \quad \ldots(1)$$

Since ϕ_2 is a Ch. Function, the bracketed expression in the above integral is non-negative. It follows that the R.H.S. in (1) is non-negative. Thus, $\sum_{j,k}\phi(t_j, t_k)\, w_j\, \bar{w}_k \geq 0$ and so $\phi(t)$ is a Ch. Function as it meets the demands of Bochner Test.

Deduction. We find that

$$\psi(t) = \frac{1}{t}\int_0^t \phi_1(u)\, du = \int_{-\infty}^{\infty}\frac{1}{t}\int_0^t e^{iux}\, du\, dF(x), = \int_{-\infty}^{\infty}\frac{1}{x}\left(\frac{e^{itx}-1}{it}\right) dF(x) \quad \text{[Def. of } \phi_1(u)\text{; integ. w.r.t. } u\text{]}$$

Recall that if X is $U(0, x)$, then $\phi_2(t : x) = \left(e^{itx} - 1\right)/ixt$. [§9-20(6)]. Hence, the above integral reduces to

$$\psi(t) = \int_{-\infty}^{\infty} \phi_2(t, x)\, dF(x).$$

By first part, this is as well a characteristic function.

9-50. Summary of Some Fundamental Results

1. ***Marcinkiewiez Theorem*** : If $\phi(t : X)$ is of the form $e^{P(t)}$, where $P(t)$ is a polynomial in t, then $P(t)$ must be of the form $P(t) = i\alpha t - h^2t^2$, α and h being real.

It follows that e^{-t^β} is a characteristic function, iff $\beta = 2$.

[Proof of this theorem is beyond the scope of this book]

2. ***Polya's (sufficient but not necessary) Conditions*** : $\phi(t)$ is a Ch. Function if

(1) $\phi(0) = 1$. (2) $\phi(t) = \phi(-t)$. (3) $\phi(t)$ is continuous. (4) $\lim_{t \to \infty} \phi(t) = 0$.

(5) $\phi(t)$ is convex for $t > 0$, i.e. for $t_1, t_2 > 0$, $\phi\left[\frac{1}{2}(t_1 + t_2)\right] \le \frac{1}{2}\left[\phi(t_1) + \phi(t_2)\right]$.

Non-necessity. If X is $N(1, 1)$, then $\phi(t : X) = e^{it - \frac{1}{2}t^2}$. But $\phi(-t) \neq \phi(t)$. This shows that the above condition are not necessary.

Note. From Polya's conditions it follows that functions such as $e^{-|t|}$ and $[1 + | t |]^{-1}$ are Ch. Functions.

3. ***Cramer's Criterion*** : A bounded and continuous function $\phi(t)$ is a Ch. Function iff

(i) $\phi(0) = 1$. (ii) $\psi(x, A) = \int_0^A \int_0^A \phi(t - u)\, e^{i(t-u)x}\, dt\, du \ge 0$.

$\psi(x, A)$ is real and non-negative, for all real x and for all $A > 0$.

4. ***Test for a function not to be a characteristic function***

Theorem : If $\phi(t)$ is not constant, and if for small t,

$$\phi(t) = 1 + g(t) + 0(t^2), \qquad (t : \text{small})$$

where $g(-t) = -g(t)$, i.e. odd function and $g(t) = O(t)$; the $\phi(t)$ is not a Ch. Function.

Proof. Let $\phi(t)$ be a Ch. Function, then

$$\psi(t) = \phi(t)\, \phi(-t) = 1 + O(t^2)$$

is also a Ch. Function. Now for any Ch. Function ϕ, with t and h real,

$$| \phi(t) - \phi(t + h) |^2 \le 2\phi(0)\, [\phi(0) - \phi_0(h)], \qquad [\phi(0) = 1] \quad [\S 9\text{-}32]$$

We apply this inequality to the function $y(t)$ and obtain

$$| \psi(t) - \psi(t + h) |^2 = O(t^2)$$

so that $\psi'(t) = 0$ everywhere, which results in a contradiction.

Remarks. $e^{-t^4}, e^{-|t|^n}\,(n > 2); (1 + t^4)^{-1}$ are not Ch. Functions.

9-60. Inversion Theorem

If $\phi_X(t)$ is the Ch. Function corresponding to c.d.f. $F(x)$ and $(a - \delta, a + \delta)$ is a continuity interval of $F(x)$ then

$$F(a + \delta) - F(a - \delta) = \lim_{T \to \infty} \int_{-T}^{T} \frac{\sin \delta t}{t} \cdot e^{-ita}\, \phi(t)\, dt. \qquad \ldots(1)$$

Proof. Consider the integral

$$J = \frac{1}{\pi}\int_{-T}^{T} \frac{\sin \delta t}{t} \,.\, e^{-ita}\phi(t)\,dt = \frac{1}{\pi}\int_{-T}^{T} \frac{\sin \delta t}{t} \,.\, e^{-ita} \int e^{itx}\,dF(x)\,dt.$$

Since $|(\sin \delta t/t)\, e^{-it(a-x)}| < \delta$, the order of integration can be interchanged so that J can be written as

$$J = \frac{1}{\pi}\int_{-\infty}^{\infty} dF(x) \int_{-T}^{T} \frac{\sin \delta t}{t} e^{-it(a-x)}\,dt = \int_{-\infty}^{\infty} \psi(T,x)\,dF(x) \qquad \ldots(2)$$

where $\psi(T,x) = \frac{1}{\pi}\int_{-T}^{T} \frac{\sin \delta t}{t} e^{-it(a-x)}\,dt = \frac{1}{\pi}\int_{-T}^{T} \frac{\sin \delta t}{t} \cos(x-a)\,t\,dt.$

Note that the second term : $(\sin \delta t/t) \sin (a-x)\,t$, in the integrand, being an odd function of t, its integral over $[-T, T]$ is zero. We now use : $2 \sin A \cos B = \sin (A+B) + \sin (A-B)$, and above gives

$$\psi(T,x) = \frac{1}{\pi}\int_{0}^{T} \frac{\sin (x-a+\delta)t}{t}\,dt + \frac{1}{\pi}\int_{0}^{T} \frac{\sin (\delta+a-x)t}{t}\,dt \qquad \ldots(3)$$

Now recall the well known result :

$$\frac{1}{\pi}\int_{0}^{\infty} \frac{\sin \lambda t}{t}\,dt = \begin{cases} \frac{1}{2}, & \text{if } \lambda > 0 \\ 0, & \text{if } \lambda = 0 \\ -\frac{1}{2}, & \text{if } \lambda < 0 \end{cases} \qquad \ldots(4)$$

Now letting $T \to \infty$ in (3) and using (4) we obtain

$$\lim_{T\to\infty} (T,x) = \begin{cases} 0, & \text{for } x < a-\delta \text{ or } x > a+\delta \\ \frac{1}{2}, & \text{for } x = a-\delta \text{ or } x = a+\delta \\ 1, & \text{for } a-\delta < x < a+\delta. \end{cases}$$

Observe that : In the first case, $\psi \to -\frac{1}{2}+\frac{1}{2}$ or $\frac{1}{2}+(-\frac{1}{2})$; in the second case, $\psi \to 0+\frac{1}{2}$ or $\frac{1}{2}+0$; in the third case $\psi \to \frac{1}{2}+\frac{1}{2}$.

Furthermore, $|\psi(T,x)|$ is bounded, since $\int_0^T [(\sin \lambda t)/t]dt$ is uniformly bounded in T. Hence by usual Lebesgue dominated convergence theorem, we find that (1) yields

$$\lim_{T\to\infty} \int_{-\infty}^{\infty} \psi(T,x)\,dF(x) = \int_{-\infty}^{\infty} \lim_{T\to\infty} \psi(T,x)\,dF(x) = \int_{a-\delta}^{a+\delta} dF(x) = F(a+\delta) - F(a-\delta)$$

Remark. If there are two distribution functions, they agree at all continuity points of both the distributions and hence are identical.

Cor. If $|\phi(t)|$ is integrable, then from (1),

$$\lim_{\delta\to 0} \frac{F(a+\delta)-F(a-\delta)}{2\delta} = \frac{1}{2\pi}\int_{-\infty}^{\infty} \lim_{\delta\to 0}\left(\frac{\sin \delta t}{\delta t}\right) e^{-ita}\,\phi(t)\,dt$$

i.e. $F'(a) = \frac{1}{2}\int_{-\infty}^{\infty} e^{-ita}\phi(t)\,dt$. Equivalently : $f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx}\phi(t)\,dt.$

9-61. Inversion Formula

If $\phi(t)$ is the Ch. Function of a variate X, then the density f of X is given by

$$\boxed{f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx}\phi(t)\,dt} \qquad ...(1)$$

Proof. We give a formal proof based on the properties of Dirac delta function :

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx}\,dt = \delta(x); \int_{-\infty}^{\infty} f(y)\,\delta(x-y)\,dy = f(x). \quad \text{[Shifting property]} \quad ...(2)$$

We consider the r.h.s. of (1) and insert the definition $\phi(t) = \int e^{ity} f(y)\,dy$ in it. This gives

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx}\,\phi(t)\,dt = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx}\left[\int_{-\infty}^{\infty} e^{ity} f(y)\,dy\right]dt \text{ [Change Integ. order] [by } \delta\text{-properties (2)]}$$

$$= \int_{-\infty}^{\infty} f(y)\left[\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-it(x-y)}\,dt\right]dy, = \int_{-\infty}^{\infty} f(y)\,\delta(x-u)\,dy = f(x)$$

Another Proof. A standard integral from Advanced Calculus is

$$\frac{1}{\pi}\int_{-\infty}^{\infty} \frac{\sin at}{t}\,dt = \begin{cases} -1, & a<0 \\ 0, & a=0. \\ 1, & a>0 \end{cases}$$

Consider a function $g(t) = (1 - \cos at)/t$. It is an odd function of t, so integral of $g(t)$ over a symmetric range $]-\infty, \infty[$ is zero. We add this to the above result to get

$$\frac{1}{\pi}\int_{-\infty}^{\infty} \frac{\sin at + i(1-\cos at)}{t}\,dt = \begin{cases} -1, & a<0 \\ 0, & a=0 \\ 1, & a>0. \end{cases}$$

Now use $e^{iat} = \cos at + i \sin at$, put $a = X - x$, multiply the above integral by 1/2, then add 1/2 to both sides, this gives

$$Y = \frac{1}{2} - \frac{i}{1\pi}\int_{-\infty}^{\infty} \frac{1-e^{it(X-x)t}}{t}\,dt = \begin{cases} 1, & X<x \\ \frac{1}{2}, & X=x \\ 0, & X>x \end{cases} \qquad ...(1)$$

For a fixed value of x, the Eq. (1) is a function of the variate X, hence we treat it to define a new random variable Y. Obviously,

$Y = 1$ with $P(X < x)$; $Y = 1/2$ with $P(X = x)$; $Y = 0$ with $P(X < x)$.

$$\therefore \qquad \text{E}(Y) = 1P(X<x) + \tfrac{1}{2}P(X=x) + 0P(X>x).$$

Since X is continuous, $P(X = x) = 0$ if x is a point of continuity in the distribution of X.

$$\therefore \qquad \text{E}(Y) = P(X<x) = F_X(x). \qquad ...(2)$$

Taking expected value of the first two members in (1) we get

$$\text{E}(Y) = \frac{1}{2} - \frac{i}{2\pi}\int_{-\infty}^{\infty} \frac{1-\text{E}[e^{i(X-x)t}]}{t}\,dt = \frac{1}{2} - \frac{i}{2\pi}\int_{-\infty}^{\infty} \frac{1-e^{-ixt}\,\phi_X(t)}{t}\,dt \qquad ...(3)$$

Eliminating $\mathsf{E}(Y)$ between Eqns. (2) and (3) ; then taking derivatives, we get

$$f_X(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx}\,\phi_X(t)\,dt, \quad [\because F'(x) = f(x)] \qquad \ldots(4)$$

Equation (4) is the inversion formula (or Fourier Transform) for a Continuous Distribution.

Comments. Inversion formula is applicable if we know that $\phi(t)$ comes from a continuous r.v. which may not be true. There is no nec. and suff. condition on ϕ for it to be from a continuous r.v. A sufficient condition is that $\int_{-\infty}^{\infty} |\phi(t)|\,dt < \infty$ (i.e. finite). However, this condition holds for Normal distribution, but not for Expo (λ). We are stuck !

Inversion in case of a discrete variate : A standard integral formula is

$$\frac{1}{2k}\int_{-k}^{k} e^{iat}\,dt = \begin{cases}(\sin ak)/ak, & a \neq 0\\ 1, & a = 0\end{cases}.$$

Take $a = X - x$, proceed to limit as $k \to \infty$ to get a new variate Y defined by

$$Y = \lim_{k\to\infty}\frac{1}{2k}\int_{-k}^{k} e^{i(X-x)t}\,dt = \begin{cases}0 & X \neq x\\ 1 & X = x\end{cases}.$$

So :
$$\mathsf{E}(y) = 0\,P(X \neq x) + 1\,P(X = x) = P(X = x)$$

$\therefore$
$$P_X(x) = \lim_{k\to\infty}\frac{1}{2k}\int_{-k}^{k} \mathsf{E}\left[e^{i(X-x)t}\right]dt = \lim_{k\to\infty}\frac{1}{2k}\int_{-k}^{k} e^{-itx}\,\phi_X(t)\,dt \qquad \ldots(5)$$

This is the desired Inversion Formula,

Note.
$$\phi_X(t) = \int_{-\infty}^{\infty} e^{itx} f(x)\,dx;\ f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx}\,\phi_X(t)\,dt, \quad [X : \text{continuous}]$$

$$\phi_X(t) = \Sigma_i\, e^{itx_i}\, p_X(x_i);\ p_X(x) = \lim_{k\to\infty}\frac{1}{2k}\int_{-k}^{k} e^{-itx}\,\phi_X(t)\,dt, \quad [X : \text{discrete}]$$

Remark. The p.d.f. for discrete variates can, generally, be read directly without resorting to their Ch. Functions.

9-62. *Uniqueness Theorem*

The characteristic function uniquely determines a p.d.f.

Proof. If possible, let there be two p.d.f.'s say f_1 and f_2 corresponding to the given Ch. Function $\phi(t)$. Then by Inversion theorem

$$f_1(x) = \int_{-\infty}^{\infty}\frac{e^{-itx}\,\phi(t)\,dt}{2\pi},\quad f_2(x) = \int_{-\infty}^{\infty}\frac{e^{-itx}\,\phi(t)\,dt}{2\pi}.$$

But this shows : $f_1(x) \equiv f_2(x)$; whence the result.

9-70. *Ch. Function of Cauchy Distribution*

$$\boxed{g(y) = \frac{\lambda}{\pi[\lambda^2 + (y-a)^2]},\ -\infty < y < \infty; \qquad \phi_Y(t) = e^{iat - \lambda|t|}.}$$

Recall : If $X \sim$ Lap (a, λ) then its density and Ch. Function are

$$f(x) = \frac{1}{2}\lambda\, e^{-\lambda|x-a|},\ -\infty < x < \infty; \qquad \phi(t) = \lambda^2\, e^{iat}/(\lambda^2 + t^2).$$

Also : $$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \phi(t)\, e^{-itx}\, dt \qquad \text{(Inversion Formula)}$$

We substitute the results of $f(x)$ and $\phi(t)$ simultaneously in the Inversion formula to get

$$\therefore \quad \frac{1}{2}\lambda\, e^{-\lambda|x-a|} \equiv \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{\lambda^2\, e^{iat}\, e^{-itx}\, dt}{\lambda^2+t^2} \qquad [\text{Put } a - x = \theta]$$

$$\text{i.e.} \quad e^{-\lambda|\theta|} \equiv \frac{\lambda}{\pi}\int_{-\infty}^{\infty} \frac{e^{it\theta}\, dt}{\lambda^2+t^2} = \frac{\lambda}{\pi}\int_{-\infty}^{\infty} \frac{e^{i\theta(y-a)}dy}{\lambda^2+(y-a)^2}. \qquad [y - a = t]$$

This we may rewrite as

$$e^{ia\theta-\lambda|\theta|} \equiv \int_{-\infty}^{\infty} \frac{\lambda\, e^{i\theta y}\, dy}{\pi[\lambda^2+(y-a)^2]} = \int_{-\infty}^{\infty} e^{i\theta y}\,.\, g(y)\, dy \qquad \ldots(1)$$

$$\text{where} \quad g(y) = \frac{\lambda}{\pi\,[\lambda^2+(y-a)^2]}, -\infty < y < \infty;\ \lambda > 0. \qquad \ldots(2)$$

Eq. (2) show that g is the Chy (a, λ) density and Eq. (1) gives its Ch. Function. For usual familiarity, change θ to t in (1) and thus.

$$\phi_y(t) = \exp[iat - \lambda|t|].$$

Direct Evaluation. From "Calculus of Residues", we borrow evaluation

$$\int_{-\infty}^{\infty} \frac{\cos tx\, dx}{b^2+x^2} = \frac{\pi}{b} e^{-b|t|} \qquad \ldots(1)$$

Now $f(x) = \dfrac{(b/\pi)}{b^2+(x-a)^2},\ b > 0;\ -\infty < x < \infty$

$$M_X(it) \doteq \mathrm{E}(e^{itX}) = \int_{-\infty}^{\infty} \frac{b}{\pi}\cdot\frac{e^{itx}\, dx}{b^2+(x-a)^2} = \frac{be^{ita}}{\pi}\int_{-\infty}^{\infty} \frac{e^{itz}\, dz}{b^2+z^2} \qquad [z = x - a]$$

$$= \frac{b}{\pi} e^{ita}\int_{-\infty}^{\infty} \frac{\cos tz\, dz}{b^2+z^2} \qquad \left[\int_{-\infty}^{\infty} \frac{\sin tz\, dz}{b^2+z^2} = 0,\ e^{i\theta} = \cos\theta = i\sin\theta\right]$$

$$= \frac{b}{\pi}\cdot e^{ita}\cdot\frac{\pi}{b} e^{-b|t|} = e^{ita-b|t|}$$

9-71. *Sum of Two Independent Cauchy Variates*

Let $X \sim$ Chy (a_1, b_1) and $Y \sim$ Chy (a_2, b_2) be independent. Then

$$(X + Y) \sim \text{Chy } (a_1 + a_2\ b_1 + b_2). \qquad \ldots(1)$$

Proof. Here $$\phi_X(t) = e^{ia_1t-b_1|t|},\ \phi_Y(t) = e^{ia_2t-b_2|t|}$$

$$\phi(t : X+Y) = \phi_X(t)\,.\,\phi_Y(t) = e^{i(a_1+a_2)t-(b_1+b_2)|t|}.$$

This shows that result (1) holds.

Cor. 1. If indep. $X_k \sim$ Chy (a_k, b_k), $1 \le k \le n$, then $S_n = X_1 + \ldots + X_n$ is Chy $(\Sigma a_k, \Sigma b_k)$.

Cor. 2. If $X_1, \ldots, X_n$ are i.i.d. Chy (a, b) variates, then $X_n = (S_n/n)$ is also Chy (a, b).

$$\phi\left(t:\frac{S_n}{n}\right) = \phi\left(\frac{t}{n}:S_n\right)=\left\{\phi\left(\frac{t}{n}:X_1\right)\right\}^n = e^{iat-b|t|}.$$

9-80. Joint Characteristic Function

Definition. The joint Ch. Function of two variates X and Y is given by

$$\phi_{X,Y}(t_1,t_2) = \mathsf{E}\{\exp(it_1 X + it_2 Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\, e^{it_1 x + it_2 y}\, dx\, dy.$$

where t_1 and t_2 are real numbers.

By setting $t_1 = 0$ or $t_2 = 0$, the Ch. Functions of X and Y are obtained, the so called *marginal* (individual) Ch. Functions :

$$\phi_X(t_1) = \phi(t_1, 0); \quad \phi_Y(t_2) = \phi(0, t_2).$$

The joint moments can be obtained as follows :

$$\mathsf{E}(X^r Y^s) = (-i)^{r+s} \frac{\partial^{r+s}\phi(0,0)}{\partial t_1^r\, \partial t_2^s}.$$

Using inverse Fourier transform, the inversion formula is

$$f(x,y) = \frac{1}{(2\pi)^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \phi(t_1,t_2)\, e^{-it_1 x - it_2 y}\, dt_1\, dt_2.$$

Factor Theorem. The jointly distributed variates X and Y are independent iff

$$\phi_{X,Y}(t_1,t_2) = \phi_X(t_1)\,\phi_Y(t_2), \quad \forall \quad t_1, t_2.$$

Proof is similar to factor theorem for m.g.f's.

***Example* 1.** If $\phi(t_1,t_2) = \exp(-2t_1^2 - 8t_2^2)$, show that X and Y are uncorrelated.

Solution. Recall $\mathsf{E}(X^r Y^s) = (-i)^{r+s}\, \partial_1^r\, \partial_2^s\, \phi(0,0)$; $[\partial_1 = \partial/\partial x,\ \partial_z = \partial/\partial y]$. Now

$\partial_1\phi(t_1,t_2) = -4t_1\phi(t_1,t_2), \partial_2\phi(t_1,t_2) = -16t_2\,\phi(t_1,t_2),\ \partial_1\partial_2\,\phi(t_1,t_2) = 64\,t_1 t_2\,\phi(t_1,t_2).$

$\therefore \quad \mathsf{E}(X) = (-i)\,\partial_1\,\phi(0,0) = 0,\ \mathsf{E}(Y) = (-i)\,\partial_2\,\phi(0,0) = 0,\ \mathsf{E}(XY) = (-i)^2\,\partial_1\,\partial_2\,\phi(0,0) = 0$

Since $\mathrm{Cov}(X,Y) = \mathsf{E}(XY) - \mathsf{E}(X)\,\mathsf{E}(Y) = 0$, it follows that X and Y are uncorrelated.

Note. $\phi(t_1,t_2) = \phi(t_1,0)\,.\,\phi(0,t_2)$, which shows that X and Y are ind. Gaussian variates.

***Example* 2.** Let $\phi(t_1,t_2) = [(1-2it_1)(1-2it_2)]^{-n/2}$, $[n > 0$, integer$]$

Obtain Var (X), Var (Y) and Corr (X, Y).

Solution. Firstly we carry out the partial differentiations needed for various moments.

$$\partial_1\,\phi(t_1,t_2) = in(1-2it_1)^{-(n/2)-1}\,.(1-2it_2)^{-n/2} \quad ; \quad \partial_1\,\phi(0,0) = in$$

$$\partial_2\,\phi(t_1,t_2) = in(1-2it_2)^{-(n/2)-1}\,.(1-2it_1)^{-n/2} \quad ; \quad \partial_2\,\phi(0,0) = in$$

$$\partial_{11}\,\phi(t_1,t_2) = -2n[(n/2)+1](1-2it_1)^{-(n/2)-2}(1-2it_1)^{-n/2}\,;\,\partial_{11}\,\phi(0,0) = -n(n+2)$$

$$\partial_{12}\phi(t_1,t_2) = -n^2(1-2it_1)^{-(n/2)-1}(1-2it_2)^{-(n/2)-1} \quad ; \quad \partial_{12}\phi(0,0) = -n^2$$

Recall : $\mathsf{E}(X^r Y^s) = (-i)^{r+s}\,\partial_1^r\,\partial_2^s\,\phi(0,0)$.

$\therefore$ $\quad \mathsf{E}(X) = (-i)(in) = n = \mathsf{E}(Y),\ \mathsf{E}(X^2) = n(n+2) = \mathsf{E}(Y^2), \mathsf{E}(XY) = n(n+2).$

$\therefore$ $\quad \mathrm{Var}\,(X) = \mathrm{Var}(Y) = n(n+2) - n^2 = 2n;\ \mathrm{Cor}(X,Y) = n^2 - n^2 = 0.\ \rho_{XY} = \sigma_{XY}/\sigma_X\,\sigma_Y = 0.$

Comments. Obviously : $\phi(t_1,t_2)\,\phi(t_1,0)\,.\,\phi(0,t_2)$ which means X and Y are independents and by symmetry, identically distributed. So one-variable evaluation is sufficient. However, the above evaluations are instructive.

9-81. *Worked-out Problems*

***Example* 1.** Prove that the following conditions are equivalent :

(*i*) $M_X(ic) = 1$, (ii) $P\{(c/2\pi)\,X \in Z\} = 1$, (iii) $M_X\,(it)$ is periodic with period $|c|$.

Solution. Let us put $(c/2\pi)\,X = Y$, i.e. $X = 2\pi Y/c$. Then $M(it : Y) = M\,(cit/2\pi : X)$. Now we attempt sequentially :

(i) $\Rightarrow$ (ii). Since $M(2\pi i : Y) = 1 \Rightarrow P\{Y \in \mathbb{Z}\} = 1$ [§9-40] $\Rightarrow$ $P\{(c/2\pi)\,X \in \mathbb{Z}\} = 1$.

(ii) $\Rightarrow$ (iii). If $P\{Y \in \mathbb{Z}\} = 1$, then $M\{it : Y)$ is periodic with period 2π [§9-40]

Since $M(it : X) = M(it : 2\pi Y/c) = M(2\pi\, it/c : Y)$, so this yields X is periodic with period $|\,c\,|$.

(iii) $\Rightarrow$ (i). As $M(i0\ ;\ X) = 1$, so $M(it : X)$ is periodic with period $|\,c\,|$.

***Example* 2.** Find the distribution for which Ch. Function is

$$\phi(t) = 1 - |t/a|,\ |t| \le a;\ \ \phi(t) = 0,\ \ |t| > a.$$

Solution. By Inversion Formula, the density f is

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx}\,\phi(t)\,dt = \frac{1}{2\pi}\int_{-a}^{a}(\cos tx - i\sin tx)\left[1 - \frac{|t|}{a}\right]dt$$

$$= \frac{1}{\pi}\int_0^a\left(1 - \frac{t}{a}\right)\cos tx\,dt. \qquad \text{[By even-odd properties of definite integral]}$$

$$\pi f(x) = \left[\left(1 - \frac{t}{a}\right)\left(\frac{\sin tx}{x}\right) - \left(-\frac{1}{a}\right)\left(\frac{-\cos tx}{x^2}\right)\right]_0^a \qquad \text{[Integration by parts]}$$

Thus, $f(x) = \dfrac{(1-\cos ax)}{\pi\,ax^2}$, $\ -\infty < x < \infty$.

***Example* 3.** (*a*) Show that if $\phi(t)$ is a Ch. Function, then $|\,\phi(t)\,|$ need not be a Ch. Function.

(*b*) Show that e^{-t^4} is not a Ch. Function. (*c*) Show that $\psi(t) = \exp(-t^\alpha)$ cannot be a Ch. Function unless $\alpha = 2$.

Solution. (a) Let $\phi(t) = \dfrac{1}{3}(1 + 2e^{it})$; it is Ch. Function of bin (1, 2/3). Now

$$|\phi(t)|^2 = \phi(t)\,\overline{\phi}(t) = \left(\frac{1}{9}\right)\left[5 + 2e^{it} + 2e^{-it}\right].$$

Let $\psi(t) = |\phi(t)| = pe^{ita} + qe^{itb}$, say $(p + q = 1)$. Then

$$[\psi(t)]^2 = |\phi(t)|^2 \Rightarrow p^2e^{2tia} + q^2e^{2tib} + 2pq\, e^{ti(a+b)} = \left(\tfrac{1}{9}\right)(5 + 2e^{it} + 2e^{-it}).$$

This result is true for all t; take $a + b = 0$, $a = -b = 1/2$; then $2pq = 5/9$, $p^2 = q^2 = 2/9$. Obviously, these equations are incompatible; we infer that $|\phi(t)|$ is not a Ch. Function.

(*b*) Assume that $\psi_X(t) = e^{-t^4}$ is a Ch. Function. By analytic expansion

$$\psi_X(t) = 1 - t^4 + (t^8/2!) + \ldots \Rightarrow \mathsf{E}|X| = 0 = \mathsf{E}(X^2) \text{ so that Var } (X) = 0$$

Thus X is degenerate with $P\{X = 0\} = 1$. As such $\phi_X(t) = \mathsf{E}(e^{itX}) = 1$.

But then $\phi(t) \neq \psi(t)$ as $e^{-t^4} \neq 1, (t \neq 0)$.

(*c*) If α is complex, $\alpha = re^{i\theta}$, then $\overline{\psi}(t) = \exp(-t^{re^{-i\theta}})$; $\psi(-t) = \exp(-(t)^{re^{i\theta}})$, so $\psi(-t) = \overline{\psi}(\bar{t})$. Thus, if $\psi(t)$ is a Ch. Function, a has to be real, further

$$\overline{\psi(t)} = \psi(-t) \Rightarrow \exp(-t^{\alpha}) = \exp((-1)^{\alpha+1}t^{\alpha}) \Rightarrow (-1)^{\alpha} = 1, \text{so } \alpha = 2k,\ k = 0, \pm 1, \pm 2, \ldots$$

However, $\alpha = 0$ yields $\psi(t) = \exp(-1) =$ const. and there is no distribution with $\psi(t) =$ const.

For $\alpha = -2k\ (k > 0), \psi(t) = e^{-(1/t)^{2k}}$, so as $t \to 0$ we get $\psi(0) = e^{-\infty} = 0$. This contradicts $\psi(0) = 1$. So α cannot be an even negative integer.

For $\alpha = 2 + 2k$, $(k > 0)$ analytic expansion yields

$$\psi(t) = \exp(-t^{2+2k}) = 1 - t^{2+2k} + O(t^{2+2k}).$$

Here $\mathsf{E}(X) = 0$, $\mathsf{E}(X^2) = 0$, since terms t and t^2 are absent. Hence Var $(X) = 0$, so that X is degenerate with $\phi(t) = e^{0t} \neq \psi(t)$. Thus, α cannot be of the form $2 + 2k$. For $\alpha = 2$, $\psi(t) = e^{-t^2}$, which is Ch. Function of $N(0, 2)$.

***Example* 4.** The continuous variates X and Y have the joint Ch. Function $\phi(t_1, t_2)$. Prove that the Ch. Function of X given $Y = y$ (fixed) is

$$\psi(t_1 | y) = \int_{-\infty}^{\infty} e^{-it_2y}\,\phi(t_1, t_2)\,dt_2 \Big/ \int_{-\infty}^{\infty} e^{-it_2y}\,\phi(0, t_2)\,dt_2.$$

Solution. The p.d.f. and Ch. Function are connected by the inversions

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx}\phi(t)\,dt\,;\ \phi(t) = \int_{-\infty}^{\infty} e^{itx}\,f(x)\,dx$$

Now, we use Double-E Rule as under :

$$\phi(t_1, t_2) = \mathsf{E}(e^{it_1X}\,e^{it_2Y}) = \mathsf{E}_Y\left\{\mathsf{E}_X(e^{it_1X}\,e^{it_2Y}\,|Y)\right\} = \mathsf{E}_Y\left\{e^{it_2Y}.\mathsf{E}(e^{it_1X}\,|y)\right\}$$

$$= \mathsf{E}_Y[e^{it_2Y}.\psi(t_1|Y)] = \int_{-\infty}^{\infty} e^{it_2y}\,\psi(t_1|y)\,f_2(y)\,dy.$$

We invert it to get

$$\psi(t_1|Y)f_2(y) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-it_2y}\,\phi(t_1, t_2)\,dt_2. \qquad \ldots(1)$$

When $t_1 = 0$, $\psi(0 \mid y) \equiv 1$, so,

$$f_2(y) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-it_2 y}\,\phi(0, t_2)\,dt_2. \qquad \ldots(2)$$

Eliminating $f_2(y)$ between (1) and (2) provides the stated result.

Problems with Solutions Provided at the End of the Text

1*. Point out, which of the following are not Ch. Function :

$\phi_1(t) = 1/(1+t)$, $\phi_2(t) = \sin at$, $\phi_3(t) = \cos bt$, $\phi_4(t) = 1 - it$, $\phi_5(t) = t\,\ell n\,t$

$\phi_6(t) = 1/(1+t^2)$, $\phi_7(t) = \exp(it)$.

2*. Find the distribution of a r.v. X, if $\phi_X(t) = 1/(2e^{-it} - 1)$.

3*. Find the distribution for which Ch. Function is $\phi(t) = e^{-\lambda|t|}, \lambda > 0, -\infty < t < \infty$.

4*. If X, Y, U and V are variates with joint density

$$(2\pi)^{-2}\exp\left[-\tfrac{1}{2}(x^2+y^2+u^2+v^2)\right]; \qquad -\infty < x, y, u, v < \infty$$

show that $Z = XY + UV$ has the density $f(z) = \frac{1}{2}e^{-|z|}$, $-\infty < z < \infty$. ...(1)

5*. If X, Y, Z are independent $N(0, 1)$ variates, find the characteristic function of the pair U, V where $U = XZ$, $V = YZ$. Deduce the Ch. Function of U and prove that U and V are not independent.

6*. Find $\phi(t_1, t_2)$ when $f(x, y) = (1/2\pi)\cos(x+y)\,\text{rect}\,(x/\pi)\,\text{rect}\,[(x+y)/\pi]$.

7*. Show that it is possible that two *different* (discrete and absolutely continuous) distributions may have the same Ch. Function on $[-1, 1]$.

8*. If X is an integer-valued r.v. and $\phi(t) = E(e^{itX})$, show that

$$P\{X \equiv 0 \pmod{k}\} = \frac{1}{k}\sum_{r=0}^{k-1}\phi\left(\frac{2\pi r}{k}\right).$$

Exercises

1. (a) Show that if $\phi(t)$ is a Ch. Function, $|\phi(t)|^2$ is also a Ch. Function.

   (b) If X and Y are i.i.d. variates, show that $\phi(t : X - Y) = |\phi(t : X)|^2$.

   (c) Prove that $(\cosh \pi x)^{-1}$, $|x| < \infty$ is a p.d.f. and find its Ch. Function.

2. (a) Recognize the variate X if the characteristic function of the variate X is

$$\phi_X(t) = \left[\frac{1}{3} + \left(\frac{2}{3}\right)e^{it}\right]^6 .\exp\left[-3(1 - e^{it})\right].$$

   (b) If $\phi_X(t) = \dfrac{e^{itm}}{[1 + (\sigma^2 t^2/2)]}$, find the p.d.f. of X and evaluate $E(X)$ and Var (X), the principal distribution characteristics.

(c) If $M_X(it) = (2 + e^{it})^3/27$, identify X.

(d) Show that $[2 - M(it)]^{-1}$ is a Ch. Function.

3. Let X be an integer-valued variate and $\phi(t : X)$ its Ch. Function. Prove that

$$P(X = k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \phi\,(t : X)\, e^{-ikt}\, dt, \quad k = 0, \pm 1, \pm 2, \ldots$$

4. Let the c.d.f. of the variate X be absolutely continuous. If the p.d.f. of X is k times differentiable $(k = 0, 1, 2, \ldots)$ and if

$a = \int_{-\infty}^{\infty} |\, f^{(j)}(x)\,|\, dx$, exists for $j = 1, 2, \ldots, k$, show that $\lim\limits_{|t| \to \infty} (\,|\, t\,|^k\, |\, \phi\,(t : X)\,|\,) = 0$

5. (a) Let X be a variate such that $\mathsf{E}(X) = 0$. Show that

$$\mathsf{E}\,(\,|\, X\,|\,) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{[1 - \phi_0(t)]}{t^2}\, dt.$$

(b) Let $X \sim N(0, \theta)$ and $g(x)$ be a function such that $e^{-x} g(x) \to 0$ as $|\, x\,| \to \infty$. Using Ch. Function, or otherwise, prove that

$$\frac{d}{d\theta} \mathsf{E}\,\{g(X)\} = \frac{1}{2} \mathsf{E}\left(\frac{d^2\, g(X)}{dx^2}\right), \quad \mu_n(\theta) = \frac{n(n-1)}{2} \int_0^{\theta} \mu_{n-2}(t)\, dt.$$

6. For any real characteristic function $\phi_0(t)$, prove the inequality

$$1 + \phi_0(2t) \geq 2\,[\phi_0(t)]^2.$$

7. (a) Let X be a continuous r.v. with c.d.f. $F(x)$. Find the Ch. Functions of $Y = a + b\, F(X)$ and $Z = -\lambda\, \ln F(X)$ and recognize distributions.

(b) Prove that for $u > 0$

$$\int_{|\, x\,| \leq u^{-1}} x^2\, dF(x) \leq \frac{3}{u^2}\, [\phi(0) - \phi_0(u)], \int_{|\, x\,| \geq u^{-1}} dF(x) \leq \frac{7}{u} \int_0^u [1 - \phi_0(t)]\, dt.$$

8. Suppose that f is the density of a variate and the associated Ch. Function $\phi(t)$ is real-valued, non-negative and integrable. Show that $kf(u)$, $-\infty < u < \infty$, is the Ch. Function of a variate with density $k\phi(x)/2\pi$, where k is so chosen that $k\, f(0) = 1$.

9. (a) Let $\phi(t)$ be a Ch. Function. Show that the followings are also characteristic functions :

$$\exp\,(\lambda\,[\phi\,(t) - 1])\,;\ \phi(t) \cosh \phi(t)/\cosh 1;\ \sinh\,[\phi(t)]/\sinh 1;\ (1 - a)\,\phi(t)/[1 - a\phi(t)].$$

(b) Show that if $\phi_X(t_0) = 1$, for some $t_0 \neq 0$, then X is of lattice type assuming the values $x_n = 2n\pi/t_0$.

10. Let $\phi(t)$ be a Ch. Function of a non-negative variate X whose kth moment μ'_k is finite.

Let $\phi_k(t) = [(i)^k\, \mu'_k]^{-1}\, \phi^{(k)}\,(t)$

where $\phi^{(k)}\,(t)$ is the kth derivative of $\phi(t)$. Prove that $\phi_k(t)$ is also a Ch. Function.

11. (a) Show that if $P_X(A) = \frac{1}{2} \mathbb{I}_A(0) + \frac{1}{4} \int_{-1}^{1} \mathbb{I}_A(x)\, dx$, then

$$M_X(it) = (1/2) + (\sin 2t).$$

(b) (X, Y) is a 2-dim. variate with the possible values 0 and 1 for both X and Y with joint probabilities as shown. Find the Ch. Functions $\phi_1(t)$, $\phi_2(t)$ and $\phi(t_1, t_2)$ for X, Y, and (X, Y) respectively. Show that

$x \downarrow$ \ $y \rightarrow$	0	1
0	p_{00}	p_{01}
1	p_{10}	p_{11}

$$\phi(t_1, t_2) = \phi_1(t_1)\,\phi_2(t_2), \text{ iff } p_{00}\, p_{11} = p_{01}\, p_{10}.$$

Find necessary and sufficient conditions for X, Y to be (i) uncorrelated, (ii) independent.

(c) Find the characteristic function for the distribution

$$f(1, 1) = f(1, -1) = 1/6, f(-1, 1) = f(-1, -1) = 1/3.$$

12. Prove that the real part of a Ch. Function is again a Ch. Function. Prove further that, if $\phi_1(t) = a_1(t) + ib_1(t)$, and $\phi_2(t) = a_2(t) + ib_2(t)$, are Ch. Functions, then $a_1(t)\,a_2(t) - b_1(t)\,b_2(t)$ is a Ch. Function.

13. If X and Y are non-degenerate variates and X/Y and X are independent, prove that X/Y and Y can never be independent.

14. Show that, if $x^{(n)} \equiv x(x-1)(x-2)\ldots(x-n+1)$, then $e^{itx} = \sum_{n=0}^{\infty} \dfrac{(e^{it}-1)^2\, x^{(n)}}{n!}$.

Deduce that $\mu'_{(r)} = [D^r\, \phi(t)]_{t=0}$, where $D = d/d(e^{it})$.

15. Find the distributions, if they exist, for which the Ch. Functions are given by

(a) $(q + pe^{it})^n$, (b) $e^{-t^2\sigma^2/2}$, (c) e^{it}, (d) $\dfrac{1}{3}\cos 2t + \dfrac{2}{3}\cos t$.

16. Given $M_X(it) = 2/(3e^{it} - 1)$, $M_Y(it) = \dfrac{1}{2}e^{-it} + \dfrac{1}{3} + \dfrac{1}{6}e^{2it}$, find the distributions of variates X and Y.

17. If $Z = X_1 + X_2 + \ldots + X_n$, where X_i's are indep. variates with Ch. Function $\phi_1(t)$ and n is a variate taking on the values 1, 2, 3, ... with Ch. Function $\phi_2(t)$, show that $\phi(t : Z) = \phi_2[-i \ln \phi_1(t)]$.

18. Let $\phi_1(t) = \Sigma a_k \cos kt$, $\phi_2(t) = \Sigma a_k\, e^{ikt}$, $0 \le k \le 1$, where $a_k \ge 0$ and $\Sigma a_k = 1$. Show that ϕ_1 and ϕ_2 are Ch. Functions and determine corresponding densities.

19. Suppose $\phi(t)$ and $F(x)$ are Ch. Function and c.d.f. of r.v. X. Show that

$$\lim_{c \to \infty} J_c = F(x+0) - F(x-0), \text{ where } J_c = \frac{1}{2c}\int_{-x}^{c} e^{-itx}\, \phi(t)\, dt.$$

20. For a continuous r.v. X with Ch. Function $\phi(t)$, show that

$$F(x) - F(0) = \frac{1}{2\pi} \lim_{T \to \infty} \int_{-T}^{T} \phi(t)\, \frac{1 - e^{ixt}}{it}\, dt.$$

Deduce that two continuous variates with the same Ch. Function, must have same c.d.f.

If you try you risk failure but if you don't, failure is certain.

Nature is not human hearted. But nature never breaks her own laws.

(Leonardo Da Vinci)

Some Probability Inequalities 10

To compute exactly $P\{X \in S\}$ is not always easy. However simple bounds on these probabilities are often sufficient for the job at hand. We thus consider such bounds.

10-10. *Biename-Chebyshev's Inequality*

Let X be a r.v. with $\mathrm{E}(X^r) < \infty$. Then for any constants b and $c > 0$,

$$\boxed{P\{|X-b| \geq c\} \leq \frac{\mathrm{E}|X-b|^r}{c^r}, \text{ i.e. } P\{|X-b|<c\} \geq 1-\frac{\mathrm{E}|X-b|^r}{c^r}} \quad \ldots(1)$$

Proof. Let X be continuous with p.d.f. $f(x)$; the discrete case is very similar. We define the events

$$A = \{x : |x-b| \geq c\}, \quad B = \{x : |x-b| < c\}.$$

$$\therefore \quad \mathrm{E}|X-b|^r = \int_{-\infty}^{\infty} f(x)|x-b|^r \, dx \qquad \{]-\infty, \infty[= A \uplus B\}$$

$$= \int_A f(x)|x-b|^r \, dx + \int_B f(x)|x-b|^r \, dx \geq \int_A f(x)|x-b|^r \, dx \quad \text{[Part of above areas]} \ldots(\text{i})$$

$$\geq \int_A f(x) c^r \, dx \qquad [\text{on } A, |x-b| \geq c \Rightarrow |x-b|^r \geq c^r]$$

$$= c^r \int_A f(x) \, dx = c^r . P(A) = c^r P\{|X-b| \geq c\}$$

This yields : $P\{|X-b| \geq c\} \leq [\mathrm{E}(X-b)^r]/c^r$.

By Negation Rule : $P\{|X-b|<c\} = 1 - P\{|X-b| \geq c\}$ and thus the other result follows.

(i) ***Markov Inequality*** : $P\{|X| \geq c\} \leq \mathrm{E}|X|/c$.

Take $b = 0$, $r = 1$.

(ii) ***Chebyshev's Inequality*** : $P\{|X-\mu| \geq c\} \leq \dfrac{\mathrm{Var}(X)}{c^2}$, i.e. $P\{|X-\mu| < c\} \geq 1 - \dfrac{\mathrm{Var}(X)}{c^2}$

Take $b = \mathrm{E}(X) = \mu$, $r = 2$.

Note. A generalized version of the above inequality occurs in §10-12.

A note on Method : We generally record expected-values and then drop (truncate) some positive terms to get the inequalities of our interest. See Eq. (i) above. Mark the following problem for this method.

***Example* 1.** For the integer-valued variate X, let $P(X = k) = p_k$ be *non-increasing* in $k = 1, 2, 3, \ldots$ Show that $p_k \le 2\mu/k^2$, where $\mu = E(X)$.

Solution. Observe : $p_j \ge p_k$ for $j \le k$. We use positive-terms truncation :

$$E(X) = \sum_{j=1}^{k} jp_j + \sum_{j=k+1}^{k} jp_j \ge \sum_{j=1}^{k} jp_j \ge \sum_{j=1}^{k} jp_k = p_k \sum_{j=1}^{k} j = p_k \frac{k(k+1)}{2} \ge \frac{1}{2} k^2 p_k .$$

Thus $E(X) \ge \frac{1}{2} p_k k^2 \Rightarrow p_k \le 2\mu / k^2$.

***Example* 2.** If X and Y are i.i.d. variates, then $P\{ |X - Y| \ge t\} \le 2P\{ |X| > \frac{1}{2}t\}, \quad \forall t > 0$.

Solution. Observe : $\{|X - Y| \ge t\} \subseteq \{|X| > \frac{1}{2}t\} \uplus \{|Y| > \frac{1}{2}t\}$

$\therefore \quad P\{|X - Y| \ge t\} \le P\{|X| > \frac{1}{2}t\} + P\{|Y| > \frac{1}{2}t\} = 2P\{|X > \frac{1}{2}t\}$.

10-11. *One-Sided Chebyshev's Inequality*

(i) If $Y \sim (0, \sigma^2)$, then for any $a > 0$ and $\sigma^2 < \infty$.

$$P\{Y \ge a\} \le \sigma^2 / (a^2 + \sigma^2) \quad (Y > 0) \qquad \ldots(1)$$

(ii) However, if $X \sim (\mu, \sigma^2)$, then for any $a > 0$, $\sigma^2 < \infty$

$$P\{X \ge \mu + a\} \le \frac{\sigma^2}{a^2 + \sigma^2}, \; P\{X \le \mu - a\} \le \frac{\sigma^2}{a^2 + \sigma^2}. \qquad \ldots(2)$$

Proof. We shall use Markov inequality : $P\{Z > b\} \le E(Z)/b, \quad [Z > 0]$

As $\{Y \ge a\} = \{[Y + (\sigma^2/a)] \ge [a + (\sigma^2/a)\}$, adding a constant, so

(i) $$P\{Y \ge a\} = P\left\{Y + (\sigma^2/a) \ge a + (\sigma^2/a)\right\} \le \frac{E[Y + (\sigma^2/a)]}{a + (\sigma^2/a)}$$

Thus $$P\{Y \ge a\} \le \frac{\sigma^2/a}{(a^2 + \sigma^2)/a} = \frac{\sigma^2}{a^2 + \sigma^2} \qquad \left[E\left(Y + \frac{\sigma^2}{a}\right) = \frac{\sigma^2}{a}, \; \mu_Y = 0\right]$$

(ii) Let $Y = X - \mu$ or $Y = \mu - X$; then $E(Y) = 0$, $\sigma_Y^2 = \sigma_X^2 = \sigma^2$

$$\therefore \quad P\left\{X - \mu \ge a\right\} = P(Y \ge a) \le \sigma^2 / (a^2 + \sigma^2)$$

Thus $$P\{X \ge a + \mu\} \le \sigma^2 / (a^2 + \sigma^2)$$

$$P\{(\mu - X) \ge a\} = P\{Y \ge a\} \le \sigma^2 / (a^2 + \sigma^2)$$

$$\therefore \quad P\{X \le \mu - a\} \le \sigma^2 / (a^2 + \sigma^2).$$

Significance of Chebyshev's Inequality

1. Chebyshev's inequality is valid without any assumption relating to the form of the distribution of X, the only requirement is the existence of Var (X). Consequently, it is applicable to a wide variety of distributions.

2. If the distribution of X is known, a better bound can usually be determined. For known distributions, the calculated $P(|X - \mu| \ge c)$ may be much lesser than the upper bound supplied by Chebyshev's inequality. However, there are distributions for which

the upper bound is actually attained. Thus, the upper bounds, *in general*, may not be improved upon.

3. Chebyshev's inequality is not only valid for absolutely continuous and discrete distributions, but also for such variates which do not fall in these two special categories, provided that their variances exist.

4. Since $P\{|X - \mu| < 3\sigma\} = P\{\mu - 3\sigma < X < \mu + 3\sigma\} \geq 8/9$, it follows that every density must concentrate most of its probability within a few (three) standard deviations of its mean. This explains importance of 3σ limits in Probability Theory.

***Markov's Inequality** (Independent Proof)*

If X is *non-negative* variate, then for any $a > 0$, $P(X \geq a) \leq [\mathrm{E}(X)/a]$.

Proof. Let X be continuous with p.d.f. 'f' Now by definition

$$\mathrm{E}(X) = \int_0^\infty xf(x)\,dx = \left(\int_0^a + \int_a^\infty\right) xf(x)\,dx \geq \int_a^\infty x f(x)\,dx \geq \int_a^\infty a f(x)\,dx$$

$$= a\int_a^\infty f(x)\,dx = aP(x \geq a).$$

Thus, $P(X \geq a) \leq [\mathrm{E}(X)/a]$. When X is discrete, summation replaces integration.

Cor. Let X be an arbitrary r.v. and b, n be arbitrary constants. Then

$$P\{|X - b| \geq c\} \leq [\mathrm{E}(X - b)^n]/c^n. \qquad \text{[Bienayme non-central Inequality].}$$

Proof. The r.v. $Y = |X - b|^n$, assumes only positive values. Applying Markov's Inequality to Y with $a = c^n$, we get

$$P\{|X - b|^n \geq c^n\} \leq \mathrm{E}(X - b)^n / c^n \Rightarrow P\{|X - b| \geq c\} \leq \mathrm{E}(X - b)^n / c^n.$$

Chebyshev inequality flows out when $b = \mu$ and $n = 2$.

Chebyshev's Law

Let $\{X_n\}$ be a sequence of r.v.'s with $\mathrm{E}(X_n) = \mu_n$ and Var $(X_n) = \sigma_n^2 < \infty$ for all n. If $\sigma_n^2 \to 0$ as $n \to \infty$, then

$$\lim P\{|X_n - \mu_n| \geq \varepsilon\} = 0, \text{ as } n \to \infty \qquad ...(1)$$

Proof. By Chebyshev's inequality : $P\{|X_n - \mu_n| \geq \varepsilon\} \leq (\sigma_n^2/\varepsilon^2), \quad \forall\ \varepsilon > 0$

Let $n \to \infty$, and $\sigma_n^2 \to 0$, then above inequality instantly reduces to (1).

10-12. *A General Function Inequality*

Let X be a r.v. and $g(X)$ be an arbitrary non-negative (integrable) function of the values of X which the r.v. X can assume. Then, provided $\mathrm{E}[g(X)]$ exists, with $c > 0$, $r = 0, 1, 2, ...$

$$\boxed{P\{g(X) \geq c\} \leq \mathrm{E}[g(X)]^r / c^r, \text{ i.e. } P\{g(X) < c\} \geq 1 - \mathrm{E}[g(X)]^r / c^r} \qquad ...(1)$$

Proof. Let $A = \{x : g(x) \geq c\}$, $B = \{x : g(x) < c\}$, so that $\{-\infty < x < \infty\} = A \uplus B$.

Suppose $f(x)$ is the p.d.f. of X. Then

$$\mathsf{E}[g(X)]^r = \int_{-\infty}^{\infty} [g(x)]^r f(x)\,dx = \int_{A \uplus B} [g(x)]^r f(x)\,dx = \int_A [g(x)]^r f(x)\,dx + \int_B [g(x)]^r f(x)\,dx$$

$$\geq \int_A [g(x)]^r f(x)\,dx \quad \text{[Part-Area of curve } [g(x)]^r f(x) \text{ on sub-interval]} \qquad \text{...(i)}$$

Now on A, $[g(x)]^r \geq (c)^r$, hence (i) reduces to

$$\mathsf{E}[g(X)]^r \geq \int_A (c)^r f(x)\,dx = (c)^r \int_A f(x)\,dx = (c)^r \,.\, P\{A\}.$$

This amounts to, with $A = \{g(X) \geq c\}$, the result

$$P\{g(X) \geq c\} \leq \mathsf{E}[g(X)]^r / c^r, \; r = 0, 1, 2, \ldots$$

The companion in (1) is negation-statement $P(\bar{A}) = 1 - P(A)$.

(i) ***Markov Inequality :*** $P\{(X) \geq c\} \leq \mathsf{E}(X)/c$.

Take $g(X) = X$, $r = 1$.

(ii) ***Chebyshev's Inequality :*** $P\{|X - \mu| \geq c\} \leq \text{Var}(X)/c^2$

Take $\qquad g(X) = |X - \mu|, r = 2, \; \text{Var}(X) = \mathsf{E}(X-\mu)^2$.

(iii) ***Bienayme-Chebyshev's Inequality :*** $P\{|X - b| \geq c\} \leq \mathsf{E}\,|X - b|^r / c^r$

Cor. Generalized Bienayme-Chebyshev's Inequality :

$$P\{|X| \geq k\} \leq \mathsf{E}[g(X)]^r / [g(k)]^r, \quad r = 0, 1, 2, \ldots \qquad \text{...(2)}$$

Proof. The non-negative non-decreasing function g, attains the value $c = g(k)$ at $x = k$ and $g(X) \geq g(k) \Leftrightarrow X \geq k$ where $X > 0$. Thus (1) yields. (2) instantly.

Note. (i) $g(X) = e^{tX} \Rightarrow P\{X \geq a\} \leq M(t)/e^{at}$

(ii) $g(X) = X^2 \Rightarrow P\{X \geq a\} \leq \mathsf{E}(X^2)/a^2 \Rightarrow P\{|X - \mu| \geq a\} \leq \sigma^2/a$.

Comments. The virtue of Chebyshev's inequality is its generality (distribution free property) and not its sharpness. Chebyshev's inequality supports the interpretation of the variance as a measure of the tendency of a variate to deviate from its mean.

Case of Var (X) = 0. In the extreme case, when $\sigma^2 = 0$, then $P(X = \mu) = 1$.
If $\sigma^2 = 0$, then $P\{|X - \mu| < c\} \geq 1 - \sigma^2/c^2 = 1$, $\forall\, c > 0$. Take $c = 1/n$ and let $n \to \infty$.

Then $c \to 0$, and using continuity of P we have $P\{|X - \mu| = 0\} = 1$.

Consult Example 3(e) §10-21.

Cor. 3. Bernstein's Inequality : $P\{X > c\} \leq e^{-ct} M_X(t), \quad t \geq 0$

Take $g(X) = e^{tX}$ and $k = e^{ct}$, identify $\{e^{tX} > e^{ct}\} = \{X > c\}$.

Deduction. Let $X_1, X_2, \ldots, X_n$ be i.i.d. variates with common m.g.f. $M(t)$ and $S_n = \sum X_i$.

Then $P\{S_n > ns\} \leq e^{-nst}[M(t)]^n$, $\forall\; s > 0$, $\forall\; t > 0$.

$M(t : S_n) = M(t : X_1 + \ldots + X_n) = [M(t : X)]^n = M^n(t)$. Take $X = S_n$, $c = ns$ in Bernsteins inequality.

Chernoff Bound : $P\{X \geq c\} \leq \min\limits_{t \geq 0} e^{-ct}\, \mathsf{E}(e^{tX})$. ...(1)

That is, $P\{X \geq c\}$ can be no larger than the smallest value of the R.H.S.

Note. Chernoff Bound (1) frequently proves tighter than the Chebyshev's inequality bound.

Observation. Let $P(X \leq 0) = 0$ and $\mathsf{E}(X) = \mu < \infty$. Show that

(i) $P(X > 2\mu) \leq 1/2$. (ii) $P(X \leq \mu t) \geq 1 - (1/t), \quad \forall \quad t \geq 1$.

(i) Let $g(X) = X$ and $k = 2\mu$. Then Chebyshev-Bienayme inequality :

$P[g(X) \geq k] \leq \mathsf{E}[g(X)]/k$ provides $P(X \geq 2\mu) \leq \mu/2\mu = 1/2$.

(ii) Take $g(X) = X$, $k = \mu t$; the above in-equality provides :

$$P(X \geq \mu t) \leq 1/t \;\Rightarrow\; P(X < \mu t) > 1 - (1/t), \;\; \forall \; t \geq 1.$$

Note. Independent proofs of (i) and (ii) are instructive. Try them.

10-13. *Cauchy-Schwarz's Inequality. (C-S Inequality)*

For random variables X, Y if $\mathsf{E}(X^2).\, \mathsf{E}(Y^2) < \infty$, then

$$\boxed{|\mathsf{E}(XY)|^2 \leq \mathsf{E}(X^2).\mathsf{E}(Y^2).} \quad ...(1)$$

Equality holds iff $Y = mX$ with probability 1, where m is some constant.

Proof. Write $\mathsf{E}(X^2) = a$, $\mathsf{E}(XY) = b$, $\mathsf{E}(Y^2) = c$. Consider

$$Q(t) = \mathsf{E}\{(Y - tX)^2\} = \mathsf{E}(t^2X^2 - 2tXY + Y^2) = t^2\mathsf{E}(X^2) - 2t\mathsf{E}(XY) + \mathsf{E}(Y^2)$$

$$= at^2 - 2bt + c.$$

Obviously, $Q(t) \geq 0$, non-negative quadratic form and hence its discriminant $\Delta = 4(b^2 - ac)$ is non-positive. Thus,

$$\Delta \leq 0 \Rightarrow b^2 \leq ac, \quad \text{i.e.} \quad |\mathsf{E}(XY)|^2 \leq \mathsf{E}(X^2)\, \mathsf{E}(Y^2).$$

Now by Chebyshevs inequality, if $\mathsf{E}(Z^2) = 0$, then

$$P\{|Z| < 1/n\} = 1 \qquad (n \geq 1)$$

Let $n \to \infty$, and use continuity property of probab. measure to get

$$1 = \lim_{n\to\infty} P\{|Z| < 1/n\} = P\left\{\lim_{n\to\infty} [|Z| < 1/n]\right\} = P(Z = 0)$$

Thus $\quad P\{Z = 0\} = 1$. with $Z = Y - tX$, we get $P\{Y = tX\} = 1$.

In this case, $\quad |\mathsf{E}(XY)|^2 = \mathsf{E}(X^2)\, .\, \mathsf{E}(Y^2)$.

Note 1. Take $Y = 1$, replace X by $|X - \mu|$; then (1) provides

$$[\mathsf{E}|X - \mu|]^2 \leq \mathsf{E}(X - \mu)^2 \;\Rightarrow\; \text{M.a.D.} < \text{S.D.}$$

Note 2. Measuring X, Y from their means. Schwarz's inequality (1) provides $\sigma_{xy} \leq \sigma_x \sigma_y$.

Cor. Triangle Inequality : $[\mathsf{E}(X+Y)^2]^{1/2} \leq \mathsf{E}[(X^2)]^{1/2} + [\mathsf{E}(Y^2)]^{1/2}$.

Proof. We observe that : $(X+Y)^2 \leq |X|(|X+Y|) + |Y|(|X+Y|)$

$$\mathsf{E}(X+Y)^2 \leq \mathsf{E}[|X|(|X+Y|)] + \mathsf{E}[|Y|(|X+Y|)] \quad ...(A)$$

Apply Schwarz inequality : $\mathsf{E}(UV) \le [\mathsf{E}(U^2)\mathsf{E}(V^2)]^{1/2}$ to each of the two terms in (A) to get

$$\mathsf{E}[(X+Y)^2] \le [\mathsf{E}(X^2)\mathsf{E}(X+Y)^2]^{1/2} + [\mathsf{E}(Y^2)\mathsf{E}(X+Y)^2]^{1/2}$$

Dividing throughout by $[\mathsf{E}(X+Y)^2]^{1/2} \ne 0$, we get the stated result.

Example : Let $X_1, X_2, ..., X_{2n}$ be i.i.d. variates and k is any real number. Set $X_1 + X_2 + ... + X_r = S_r$. Show that

$$P\left\{\left|\frac{S_{2n}}{2n} - k\right| \le \left|\frac{S_n}{n} - k\right|\right\} \ge \frac{1}{2}. \qquad ...(1)$$

Solution. Let $Y = (S_n/n)$ and $Z = (S_{2n} - S_n)/n$. If event in (1) is B (say), then

$$B = \{|(Y+Z)/2 - k| \le |Y-k|\} = \{|(Y-k)+(Z-k)| \le 2|Y-k|\}$$

$$= \{|Z-k| < |Y-k|\}$$

$$\therefore \quad P(B) = P\{|Z-k| < |Y-k| \ge (\tfrac{1}{2}) \qquad ...(2)$$

This follows from the fact that X_j are i.i.d. variates and

$$P\{X_j > X_k\} \ge (\tfrac{1}{2})$$

In fact, the inequality in (2) is true since, for discrete variates $P\{X_1 = X_2\} > 0$, and for continuous variates, $P\{X_1 = X_2\} = 0$.

10-20. *Kolmogorov's Inequality*

Let $X_1, X_2, ..., X_n$ be n *independent centred* variates and Var $(X_k) = \sigma_k^2 < \infty, 1 \le k \le n$. Write $S_k = X_1 + X_2 + ... + X_n$. Then for all $c > 0$,

$$P\left\{\max_{1\le k\le n} |S_k| > c\right\} \le \sum_{k=1}^{n} \frac{\sigma_k^2}{c^2}. \qquad ...(1)$$

Proof. Introduce the r.v. N by the following definition :

$N = \min_k \{k \le n\}$ if $S_k^2 > c^2$; $N = n$ if $S_k^2 \le c^2, \ \forall \ \ 1 \le k \le n$. Explicity :

$N = 1$ if $X_1^2 > c^2$, $N = 2$ if $X_1^2 \le c^2$ and $S_2^2 > c^2$, ...

$N = k$ if $X_1^2 \le c^2, ..., S_{k-1}^2 \le c^2, S_k^2 > c^2, ...,$ $\qquad N = n$ if $S_k^2 \le c^2, \ \ 1 \le k \le n-1$.

We apply Markov's inequality to the identical events $\left\{\max_{1\le k\le n} (S_k^2) > c^2\right\} \equiv \{S_N^2 > c^2\}$ to get

$$P\{\max S_k^2 > c^2\} = P\{S_N^2 > c^2\} \le \mathsf{E}(S_N^2)/c^2. \qquad ...(2)$$

Note that $\qquad \mathsf{E}(S_n)^2 = \text{Var } S_n = \sum_{k=1}^{n} \text{Var}(X_k) = \sum_{k=1}^{n} \sigma_k^2 \qquad ...(3)$

We find that (1) flows from (2) if $\mathsf{E}(S_N^2) \le \mathsf{E}(S_n^2)$. For this purpose, we use conditioning on N.

$$\mathsf{E}(S_n^2 \mid N=n) = \mathsf{E}(S_N^2 \mid N=n), \text{ and for } k<n,$$

$$\mathsf{E}(S_n^2 \mid N=k) = \mathsf{E}([S_k + S_{n-k}]^2 \mid N=n) = \mathsf{E}(S_k^2 \mid N=k) + 2\mathsf{E}\{(S_k . S_{n-k}) \mid N=k\} + \mathsf{E}(S_{n-k}^2 \mid N=k) \quad \ldots(4)$$

Now, the event $\{N = k\}$ contain information about $X_1, \ldots, X_k$, since the event implies that $X_1^2 \le c^2, \ldots, S_{k-1}^2 \le c^2, S_k^2 > c^2$. However, this event says nothing about the values of $X_{k+1}, X_{k+2}, \ldots, X_n$. Since all the variates are independent, it follows that S_k and $S_{n-k} = X_{k+1} + \ldots + X_n$ remain conditionally independent given $\{N = k\}$. Hence

$$\mathsf{E}\{(S_k . S_{n-k})\, N=k\} = \mathsf{E}\{\mid S_k \mid N=k\} . \mathsf{E}\{S_{n-k} \mid N=k\} = \mathsf{E}(S_k \mid N=k)\, \mathsf{E}(S_{n-k}) = 0$$

since $\mathsf{E}(S_{n-k}) = \mathsf{E}(X_{k+1}) + \ldots + \mathsf{E}(X_n) = 0$, (centred variates). Thus (4) reduces to

$$\mathsf{E}(S_n^2 \mid N=k) \ge \mathsf{E}(S_k^2 \mid N=k) = \mathsf{E}\{(S_N^2) \mid N=k\}$$

It follows that for all values of N, $\mathsf{E}(S_N^2 \mid N) \ge \mathsf{E}(S_N^2 \mid N)$. By Double-E Rule

$$\mathsf{EE}(S_n^2 \mid N) \ge \mathsf{EE}(S_N^2 \mid N) \Rightarrow \mathsf{E}(S_n^2) \ge \mathsf{E}(S_N^2), \text{i.e. } \mathsf{E}(S_N^2) \le \mathsf{E}(S_n^2) = \Sigma\sigma_k^2 . \quad \ldots(5)$$

Substituting (5) in (2) the result (1) follows at once.

Comments. (i) Take $n = 1$, then $P\{\mid X - \mu \mid > c\} \le \sigma^2/c^2$. [Chebyshev's inequality] ...(6)

(ii) Kolmogorov's inequality is much stronger than Chebyshev's inequality. For

$$P\{\mid S_n \mid > c\} \le \sum_{k=1}^{n} \frac{\sigma_k^2}{c^2} \text{ (Chebyshev's inequality)}, \quad P\left\{\bigcup_{k=1}^{n} \mid S_k \mid > c\right\} \le \sum_{k=1}^{n} \frac{\sigma_k^2}{c^2} \text{ (Kolmogorov inequality)}.$$

The set $\bigcup\{\mid S_k \mid > c\}$ is much larger than the set $\{\mid S_n \mid > c\}$.

10-21. *Worked-out Problems*

***Example* 1.** Two dice are thrown once. If X is the sum of the numbers showing up, prove that $P\{\mid X - 7 \mid \ge 3\} \le 35/54$, and compare this value with the exact probability.

Solution. If X_1, X_2 are face values of two dice, then

$$\mathsf{E}(X_1) = (1/6)(1+2+\ldots+6) = 7/2; \quad \mathsf{E}(X_1^2) = (1/6)(1^2+2^2+\ldots+6^2) = 91/6,$$

$$\text{Var}(X_1) = (91/6) - (49/4) = 35/12$$

Now $\quad X = X_1 + X_2, \mathsf{E}(X) = 2\mathsf{E}(X_1) = 7, \ \text{Var}(X) = 2\,\text{Var}(X_1) = 35/6.$

From Chebyshev's inequality $P\{\mid X - \mu \mid \ge c\} \le (\sigma^2/c^2)$, we get

$$P\{\mid X-7 \mid \ge 3\} \le 35/54, \qquad [\sigma^2/c^2 = 35/54, c=3];$$

Exact evaluation : Now let $p = P\{\mid X - 7 \mid < 3\} = P\{4 < X < 10\} = p_5 + p_6 + \ldots + p_9$ ...(1)

$$p_k = P(S_k) = \{6 - \mid k - 7 \mid\}/36, \ k = 2, 3, \ldots, 12$$

$$p_5 = 4/36,\ p_6 = 5/36,\ p_7 = 6/36,\ p_8 = 5/36,\ p_9 = 4/36.$$

Hence $p = 2/3$, $q = 1/3$ [from (1)]. Thus $q = 1/3 = P\{\,|X - 7| \geq 3\}$ is the actual value, a wide-apart result.

***Example* 2.** Obtain *Chebyshev's Inequality* in the form : $p = P\{|\overline{X}_n - \mu| < \varepsilon\} \geq 1 - \delta$.

(a) How large a sample must be taken in order that you are 99% certain that $\overline{X}_n$ is within $\sigma/2$ of μ ?

(*b*) How large a sample must be taken in order that the probability will be at least 0.95 that $\overline{X}_n$ will lie within 0.5 of μ ? μ is unknown and $\sigma = 1$.

(*c*) How large n is to be so that $P\{|\overline{X}_n - \mu| < 1\} > 0.9$, when $\sigma^2 = 1$? [$n = 10$]

Solution. By Chebyshev's inequality : $p = P\{|\overline{X}_n - \mu| < \varepsilon\} \geq 1 - [\mathsf{E}(\overline{X}_n - \mu)^2/\varepsilon^2] = 1 - (\sigma^2/n\varepsilon^2)$

$\therefore \quad p \geq 1 - (\sigma^2/n\varepsilon^2) \geq 1 - \delta$ if $\delta = \sigma^2/n\varepsilon^2$ or $n = \sigma^2/\delta\varepsilon^2$.

Note. If we want p to equal at least λ, we must put $\lambda = 1 - (\sigma^2/n\varepsilon^2)$, from which we can solve for n, giving $n = \sigma^2/(1-\lambda)\varepsilon^2$.

(a) Here $\varepsilon = \sigma/2$, $\lambda = 0.99$. Thus $n = 400$. (*b*) Here $\lambda = 0.95$, $\sigma^2 = 1$, $\varepsilon = 0.5$, $n = 80$.

***Example* 3. Total-E Rule.** Define $\mathsf{E}(X \mid A)$. Let $A_1, A_2, \ldots, A_n$ be a partition of S and B an arbitrary event of S. Show that

$$\mathsf{E}(X) = P(A_1)\mathsf{E}(X \mid A_1) + P(A_2)\mathsf{E}(X \mid A_2) + \ldots + P(A_n)\mathsf{E}(X \mid A_n) \qquad \ldots(1)$$

Deduce the following :

(a) $\quad P(B) = P(A_1)\,P(B \mid A_1) + P(A_2)\,P(B \mid A_2) + \ldots + P(A_n)\,P(B \mid A_n)$. $\qquad \ldots(2)$

(*b*) $$P\{|X - c| > a\} \leq \mathsf{E}(|X - c|^r)/a^r. \qquad \ldots(3)$$

(*c*) $$\text{Var}\,(X) = 0, \text{ iff } P\{X = \mu\} = 1. \qquad \ldots(4)$$

Solution. Here I_A is indicator variable of event A, I_j = indicator of event A_j, $1 \leq j \leq n$.

Definition. $\quad \mathsf{E}(X \mid A) = \mathsf{E}(XI_A)/\mathsf{E}(I_A) = \mathsf{E}(XI_A)/P(A) \qquad \ldots(5)$

$$\begin{aligned}\mathsf{E}(X) &= \mathsf{E}(XI_S) = \mathsf{E}\{X(I_1 + I_2 + \ldots + I_n)\} = \mathsf{E}(XI_1 + XI_2 + \ldots + XI_n) \\ &= \mathsf{E}(XI_1) + \mathsf{E}(XI_2) + \ldots + \mathsf{E}(XI_n) \qquad \text{[by Lin E]} \\ &= P(A_1)\mathsf{E}(X \mid A_1) + P(A_2)\mathsf{E}(X \mid A_2) + \ldots + P(A_n)\mathsf{E}(X \mid A_n). \qquad \text{[by (5)]}\end{aligned}$$

Deductions. (a) Let $X = I_B$, then (1) yields

$$P(B) = \Sigma P(A_j)\mathsf{E}(I_B \mid A_j) = \Sigma P(A_j)\,P(B \mid A_j).$$

(b) Let $X > 0$ and define $A = \{X \leq a\}$, $\overline{A} = \{X > a\}$. Now for events $A, \overline{A}$ formula (1) provides

$$\mathsf{E}(X) = \mathsf{E}(X \mid A)\,P(A) + \mathsf{E}(X \mid \overline{A})\,P(\overline{A})$$

$$\mathsf{E}(X^r) = \mathsf{E}(X^r \mid X \le a)\, P(X \le a) + \mathsf{E}(X^r \mid X > a)\, P(X > a) \ge \mathsf{E}(X^r \mid X > a)\, P(X > a).$$

Since $\mathsf{E}(X^r \mid X > a) > a^r$, this inequality yields

$$\mathsf{E}(X^r) > a^r\, P(X > a) \;\Rightarrow\; P\{X > a\} < \mathsf{E}(X^r)/a^r \qquad \text{[Markov Inequality]}$$

Replace X by $\mid X - c \mid$ to get result (3).

Note. Take $c = \mathsf{E}(X) = \mu$, $r = 2$ to get result $P\{\mid X - \mu \mid > a\} < \sigma^2 / a^2$ [Chebyshev's Inequality]

$$\text{(c)} \qquad \{X = \mu\} = \bigcap_{n=1}^{\infty} \left\{\mu - \frac{1}{n} < X < \mu + \frac{1}{n}\right\} = \bigcap_{n=1}^{\infty} \left\{\mid X - \mu \mid < \frac{1}{n}\right\}. \qquad \text{(Draw Figure)}$$

Let $D = \{X = \mu\}$ and $D_n = \{\mid X - \mu \mid < 1/n\}$. Observe that $\{D_n\}$ monotonically decreases to D. Hence by Continuity of Probability measure

$$P(D) = \lim_{n\to\infty} P(D_n)\,. \qquad \ldots(7)$$

By Chebyshev's Inequality :

$$P(D_n) = P\{\mid X - \mu \mid < 1/n\} \ge 1 - n^2\sigma^2 \;\Rightarrow\; P(D_n) = 1, \quad (\text{use } \sigma = 0),$$

Substituting in (6) we get $P(D) = 1 \Rightarrow P\{X = \mu\} = 1$, whenever $\sigma^2 = 0$.

Conversely, $\qquad \text{Var}(X) = \mathsf{E}(X - \mu)^2 = 0.1 = 0, \text{at } X = \mu$ with probability 1.

Problems with Solutions Provided at the End of the Text

1*. (i) A variate X has mean 50 and variance 100. Assess the following :
(a) $P(\mid X - 50 \mid \ge 15)$, (b) $P(X \ge 65) + P(X \le 35)$, (c) $P(\mid X - 50 \mid < 20)$,
(d) $P(30 < X < 70)$, (e) Values of t that make $P(\mid X - 50 \mid \ge t) \le 0.01$.

2*. Let $Z_1, Z_2, \ldots, Z_n$ be arbitrary but standardized variates. Show that

$$P\{\mid Z_k \mid \le t\sqrt{n}, 1 \le k \le n\} \ge 1 - (1/t^2), \quad \forall \; t > 0.$$

3*. Let $X_1, X_2, \ldots, X_n$ be i.i.d. variates with $\mathsf{E}(X_j^2) < \infty$.

Let $Z_n = [6/n(n+1)(2n+1)]\, \Sigma k^2 X_k, (1 \le k \le n)$. Show that $Z_n \xrightarrow{p} \mathsf{E}(X_1)$.

4*. If X is a variate such that $\mathsf{E}(X) = 3$, $\mathsf{E}(X^2) = 13$, show that

$$P(-2 < X < 8) \ge 21/25.$$

5*. A sample of size n is drawn from a population whose mean is 5 and S.D. is 1. Prove that $P\{\mid \overline{X} - 5 \mid < 0.001\} \ge 1 - (10^6/n)$.

6*. A discrete variate X is specified by $f(-a) = f(a) = 1/8$, $f(0) = 3/4$. Compute $P(\mid X \mid \ge 2\sigma)$ and compare it with Chebyshev's inequality bound.

7*. Evaluate $P(\mid X - \mu_X \mid \ge 2\sigma_X)$, for the discrete variate with density

$$f(X) = (1/8)\, I_{(-1)}(x) + (6/8)\, I_{(0)}(x) + (1/8)\, I_{(1)}(x).$$

8*. Does there exist a variate X for which $P\{\mu_X - 2\sigma \le X \le \mu_X + 2\sigma\} = 0.6$.

9*. In sampling without replacement, what sample size n will ensure that the sample mean differs at most from the population mean with 5% with probability greater than or equal to 90%. Assume the coefficient of variation is equal to 1/10.

10*. Box k contains one defective and k non-defective balls $k = 1, 2, \ldots$ A ball is drawn at random from each box. Let $X_k = 1$ if ball from kth box is defective and $X_k = 0$, otherwise. Let

$S_n = X_1 + X_2 + \ldots + X_n$. Show that as $n \to \infty$, $P\{\,|\,S_n/\ln n\,| - 1\,| < \varepsilon\} = 1$. ...(1)

10-22. *Gauss' or Camp-Meidell Inequality*

Suppose, X is a unimodal variate with mode x_0 and $s = (\mu - x_0)/\sigma$ is Pearson's measure of skewness. If $k > |\,s\,|$, and $X^* = (X - \mu)/\sigma$, then

$$P\{|X^*| \ge k\} \le 4(1 + S_k^2)/9(k - |s|)^2. \qquad \text{[Extreme value probability theorem]}$$

When $s = 0$, $\mu = x_0$, then $P\{\,|\,X^*\,| \ge k\} \le 4/9\,k^2$.

Gauss' Lemma. If $g(y)$ is strictly decreasing for $y > 0$, then for any $\lambda > 0$, show that

$$\lambda^2 \int_\lambda^\infty g(y)\,dy \le \frac{4}{9}\int_0^\infty y^2 g(y)\,dy. \qquad [\text{i.e. } \lambda^2 P\{Y > \lambda\} \le (4/9)\,E(Y^2)] \quad \ldots(1)$$

Proof. (i) In the special case, when $g(y) = \text{const.}$, say $g(y) = A$, for $0 < y \le c$; $g(y) = 0$, for $y \ge c$, we have $\varphi(\lambda) = \lambda^2 \int_\lambda^c g(y)\,dy = A\lambda^2(c - \lambda)$.

$$\varphi'(\lambda) = A(2c\lambda - 3\lambda^2), \quad \varphi''(\lambda) = A(2c - 6\lambda); \quad \varphi'(\lambda) = 0 \Rightarrow \lambda = 2c/3.$$

Thus, $\varphi(\lambda)$ is maximized for $\lambda = 2c/3$ and max $\varphi(\lambda) = 4Ac^3/27$. The R.H.S. in (1) also integrates to $4Ac^3/27$. Thus (1) holds.

(ii) In the general case, when $g(y) \ne$ const. we define the function

$$h(y) = g(y),\ 0 < y < \lambda + a; \quad h(y) = 0 \text{ for } y \ge \lambda + a.$$

From Bonnet's 2nd Mean value Theorem of Integral Calculus :

$$\int_a^b \psi(x)\,\varphi(x)\,dx = \varphi(a)\int_a^\xi \psi(x)\,dx \qquad [\phi(x)\downarrow;\ \xi \in (a, b)]$$

$$\therefore \quad \int_\lambda^\infty g(y)\,dy = g(\lambda)\int_\lambda^{\lambda+a} dy = a g(\lambda). \quad \ldots(2)$$

Now
$$\lambda^2 \int_\lambda^\infty h(y)\,dy \le \left(\frac{4}{9}\right)\int_\lambda^\infty y^2 h(y)\,dy \le \left(\frac{4}{9}\right)\int_0^\infty y^2\,h(y)\,dy \qquad \text{[By Part (i)]} \quad \ldots(3)$$

In (3),
$$\text{L.H.S.} = \lambda^2 \int_\lambda^{\lambda+a} g(\lambda)\,dy = a\lambda^2 g(\lambda) = \lambda^2 \int_\lambda^\infty g(y)\,dy \qquad \text{[by (2)]}$$

$$\therefore \quad \lambda^2 \int_\lambda^\infty g(y)\,dy \le \left(\frac{4}{9}\right)\int_0^\infty y^2\,h(y)\,dy \le \left(\frac{4}{9}\right)\int_0^\infty y^2 g(y)\,dy, \qquad [h(y) < g(y)].$$

Proof to Gauss' Inequality. Use this lemma in exercise No. 19 of set of exercises 10(b).

Example : Find Pearson's measure of skewness for a variate X defined by $f(x) = e^{-x}$, $0 \le x < \infty$. Use Chebyshev's inequality or otherwise to find an upper bound for $P\{ \mid X - 1 \mid \ge 5\}$.

Solution. $\mu_r' = \mathrm{E}(X^r) = \int_0^\infty x^r e^{-x} dx = \Gamma(r+1) = r!$. $\mu = \mu_1' = 1$, $\mu_2' = 2$, $\sigma^2 = 2 - 1 = 1$.

Since $f(0) = 1$, and $f(\infty) = 0$, it follows that $X = 0$ is the unique modal value. Clearly the modal value in this instance is not very satisfactory measure of the central tendency. Now

$$s_k = \frac{(\mu - x_0)}{\sigma} = \frac{(1-0)}{1} = 1.$$

By Chebyshev's Inequality : $P\{ \mid X \mid \ge b\} \le 1b^2 \Rightarrow P(\mid X - 1 \mid \ge 5) \le 1/25.$

Using extreme-value probability theorem;

$$P\{\mid X - 1 \mid \ge 5) \le [4(1+1)/9(5-1)^2] = 1/18.$$

Exercise 10(a)

1. Variate X takes the values -1, 1, 3, 5 with associated probabilities $\frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{2}$. Show $p = P\{ \mid X - 3 \mid \ge 1\} = 5/6$ directly and find an upper bound to p by using Chebyshev's inequality. **[Ans.** 16/3**]**
2. Let $\mathrm{E}(X) = 11$, Var $(X) = 2$. Use Chebyshev's inequality to estimate from below
$$P\{X = 9, 10, 11, 12, 13\}.$$
3. Variate X takes the values $-8, -6, 0, 6, 1$ with associated probability $\frac{1}{16}, \frac{2}{16}, \frac{10}{16}, \frac{2}{16}, \frac{1}{16}$. Obtain the upper limit for $P\{ \mid X - 6 \mid \ge 2\}$ and compare this value with actual probability.
4. (a) Let X have the p.d.f. $f(x) = 1/2\sqrt{3}, -\sqrt{3} \le x \le \sqrt{3}$; $f(x) = 0$, otherwise. Show that $P\{(X - \mu) \ge 3\sigma/2\} = 0.134$ and compare it with the upper bound obtained by Chebyshev's Inequality. **[Ans.** 0.444**]**
(*b*) Let $X \sim U(-a, a)$. Compare $P\{ \mid X \mid \ge 1.5\,\sigma\}$ with upper bound supplied by Chebyshev's Inequality.
(*c*) Let $f(x) = 1, 0 < x < 1$; $f(x) = 0$, otherwise. Obtain the lower bound to $P\{\mid X - \frac{1}{2} \mid \le 1/\sqrt{3}\}$ and compare it with exact value.
5. Variate X has p.d.f. $f(x) = e^{-x}, x \ge 0$. Use Chebyshev's inequality to show that $P\{ \mid X - 1 \mid > 2\} < \frac{1}{4}$ and show that the actual probability is e^{-3}.
6. Variate X has p.d.f. $f(x) = 2x/a^2, a > 0, 0 \le x \le a$. Show that $P\{ \mid X - \mu \mid \ge a/3\} = 1/9$ and compare it with upper bound.
7. If X is the number scored in a throw of a fair die, show that the Chebyshev's inequality gives $P\{ \mid X - \mu \mid > 2.5\} < 0.47$, while the actual probability is zero.
8. A discrete variate X can assume only the values $x = 1, 2, 3, \ldots$ with probability 2^{-x}. Show that Chebyshev's inequality gives $P\{ \mid X - 2 \mid > 2\} \le 1/2$, while the actual probability is 1/16.

9. (a) A variate X has the p.m.f. $f(1) = f(-1) = p$, $f(0) = 6p$. Show that there is a value of δ, s.t. $P(|X - \mu| \geq \delta) = \sigma^2/\delta^2$, so that, in general, the bound given by Chebyshev's inequality cannot be improved. **[Ans. $\delta = 1$]**

(b) Let $f(-2) = f(2) = 2/9$, $f(0) = 5/9$. Compute $P\{|X - \mu| \geq 3\sigma/2\}$ and comment on the result.

10. (a) Use Chebyshev's inequality to prove that if Var $(X) = 0$, then $P(X = \mu) = 1$.

(b) Does there exist a variate X for which $P\{\mu - 3\sigma \leq X \leq \mu + 3\sigma\} = 0.7$.

11. (a) Let $E(X) = \mu$, Var $(X) = \sigma^2$. If at least 99% of the values of X fall within k standard deviations from the mean, find k.

(b) For the Pareto law : $f(x) = 24x^{-4}$, $x \geq 2$, graph the function $\phi(\delta) = P\{|X - \mu| > \delta\}$ and compare it with the upper bound supplied by Chebyshev's inequality.

12. Show that $P\{X \geq 2\mu\} \leq \frac{1}{2}$, if $P(X \leq 0) = 0$.

13. Let $\mu = \sigma = \theta$ (say). Show that a lower bound for $P\{-3\theta < X < 5\theta\}$ is 15/16.

14. Let $M(t)$, $|t| < h$, be the m.g.f. of a variate X. Show that

$$P(X \geq a) \leq e^{-at} M(t), \quad 0 < t < h; \qquad P(X \leq a) \leq e^{-at} M(t), -h < t < 0.$$

If $M(t) = (e^t - e^{-t})/2t$, t real, deduce that $P(X \geq 1) = 0$ and $P(X \leq -1) = 0$.

15. Establish $P\{|X - \mu| \geq k\} \leq E[(X - \mu)^{2n}]/k^{2n}; \quad \mu = E(X), k \geq 0.$

16. If $E(X^2) < \infty$ and if $P(|X - \mu| < a) = p$, where a and p are given, find a lower bound to Var (X).

17. Let $X_1, X_2, \ldots, X_n$ be (possibly dependent) variates such that $E(X_j) = \mu_j$ and Var $(X_i) = \sigma_i^2$. $i = 1, 2, \ldots, n$. Prove that, for any positive numbers $d_1, d_2, \ldots, d_k$. $P\{|X_i - \mu_i| \leq d_i, 1 \leq i \leq n\} \geq 1 - \sum d_i^{-2} \sigma_i^2$.

18. $X_1, X_2, \ldots, X_n$ are i.i.d variates with $P(X_i = 1) = p$, $P(X_i = 0) = q$, $p + q = 1$, $0 \leq q \leq 1$. Let $W_n = (X_1/2) + (X_2/2^2) + (X_3/2^3) + \ldots + (X_n/2^n)$ and suppose that $W_n \to W$ as $n \to \infty$. Now prove or disprove :

(a) $P(W_n < 1/2) = 1, n < \infty$ (b) $P(W_n < 1/2) = q$

(c) $P[W_n < (1/2)^{n+1}] = q^{n+1}$.

10-30. Frechet Inequality

Let a and b be any two positive numbers and let for a r.v. X, $E(X) = \mu$ and Var $(X) = \sigma^2$. Then

$$p = P\{\mu - a\sigma < X < \mu + b\sigma\} \geq 1 - \frac{1 + [(b-a)/2]^2}{[(b+a)/2]^2}. \qquad \ldots(1)$$

Proof. It pays to use standardization. Let $Z = (X - \mu)/\sigma$. Then $E(Z) = 0$ and Var $Z = 1$. Now

$$p = P\{\mu - a\sigma < X < \mu + b\sigma\} = P\{-a < Z < b\}, \qquad \text{[Subtract } (b-a)/2, \text{ from all nos.]}$$

$$= P\{-\tfrac{1}{2}(a+b) < [Z - \tfrac{1}{2}(b-a)] < \tfrac{1}{2}(a+b)\}$$

$$= P\{|Z - \tfrac{1}{2}(b-a)| < \tfrac{1}{2}(a+b)\}$$

$$\geq -\{E[Z - \tfrac{1}{2}(b-a)]^2 / [\tfrac{1}{2}(a+b)]^2\}, \qquad \text{by Biename-Chebyshev's Inequality}$$

As $$E\{Z - \tfrac{1}{2}(b-a)\}^2 = E(Z^2) - (b-a)E(Z) + \tfrac{1}{4}(b-a)^2 = 1 + \tfrac{1}{4}(b-a)^2$$

$\therefore$ $$p \geq 1 - \left\{1 + \tfrac{1}{4}(b-a)^2\right\} / \tfrac{1}{4}(a+b)^2.$$

Comments. The standardized variate X^* is independent of the units in which X is measured.

Let $X_1 = aX + b$, then

$$X_1^* = \frac{(aX+b) - \mathsf{E}(aX+b)}{\sqrt{\mathrm{Var}\,(aX+b)}} = \frac{a\{X - \mathsf{E}(X)\}}{a(\sigma_X)} = \frac{X - \mu_X}{\sigma_X} = X^*.$$

10-31. *Strong Upper Bounds*

Suppose $g(X) = (X - \mu)^4$, $\mathsf{E}\,[g(X)] = \mu_4$, then by 10-30 (1), using $k = t^4$

$$P\{\,|X - \mu| \geq t\} \leq (\mu_4/t^4). \qquad \text{...(1)}$$

This will give a stronger upper bound (smaller in numerical value) than that supplied by Chebyshev's inequality for those cases, in which

$$(\mu_4/t^4) < (\sigma^2/t^2) \;\Rightarrow\; t > [\mu_4/\mu_2]^{1/2}. \qquad [\mu_2 = \sigma^2] \qquad \text{...(2)}$$

From (2), it follows that if a sufficiently large t is used, then (1) gives a smaller upper bound than that given by Chebyshev's inequality. This is natural, because large values of t imply a concern with large deviations of X from μ, and large deviations are better characterized by μ_4 rather than by μ_2.

Now suppose that $g(X) = [(X - \mu)^2 - \mu_2]^2$; $\mathsf{E}[g(X)] = \mu_4 - \mu_2^2$ and §10-30 (i) implies

$$P\{\,|(X - \mu)^2 - \mu_2| \geq t\} \leq (\mu_4 - \mu_2^2)/t^2. \qquad \text{...(3)}$$

Two observations emerge from (3) :

(i) Since probability measure is non-negative, $\mu_4 - \mu_2^2 \geq 0 \;\Rightarrow\; \beta_2 \geq 1$.

(ii) If $\mu_4 - \mu_2^2$ is sufficiently small, then there will be only a very small probability that $(X - \mu)^2$ will be very far from μ_2. If μ_2 is sufficiently large, then $(X - \mu)^2$ being close to μ_2, implies that X cannot be close to μ. Such a behaviour is characteristic of bimodal distributions.

The Chernoff Bound

Let $X_1, X_2, \ldots, X_n$ be a sequence of i.i.d. variates with m.g.f $M(t)$ and c.g.f. $K(t)$ i.e. $K(t) = \ln M(t)$ or $M(t) = \exp\,(K(t))$. Now form the sum

$$S_n = X_1 + X_2 + \ldots + X_n.$$

Then
$$E(e^{tS_n}) = \mathsf{E}[e^{t(X_1 + \ldots + X_n)}] = [\mathsf{E}(e^{tX})]^n = [M(t)]^n = e^{nK(t)}. \qquad \text{...(1)}$$

Let λ be a real number and define events A and B by

$$A = \{S_n \geq \lambda\} = \{e^{tS_n} \geq e^{\lambda t}\}\; t > 0, \quad B = \{S_n \leq \lambda\} = \{e^{tS_n} \geq e^{\lambda t}\},\; t < 0.$$

If I_A and I_B are the indicators of events A and B, then

$$e^{tS_n} \geq e^{\lambda t}\, I_A\; (t \geq 0), \quad e^{tS_n} \geq e^{t\lambda}\, I_B \;\; (t \leq 0) \qquad \text{...(2)}$$

We can verify (2) just by taking logarithms. Now expectations of (2) and use of (1) yield

$$P(A) \leq e^{nK(t) - t\lambda}, \; t \geq 0, \;\; P(B) \leq e^{nK(t) - \lambda t}, \; t \leq 0 \qquad \text{...(3)}$$

The value of t that minimizes, the R.H.S. of (3) is that value of t that minimizes $g(t) = nK(t) - \lambda t$. So by differentiation $g'(t) = nK'(t) - \lambda$,

$g''(t) = nK''(t) > 0$, (Var $(X) \not< 0$). So $g'(t) = 0 \Rightarrow nK'(t_0) = \lambda$. Thus (3) gives

$$\left.\begin{aligned} P\{S_n \geq nK'(t_0)\} &\leq e^{-n[t_0 K'(t_0) - K(t_0)]}, \quad t_0 \geq 0 \\ P\{S_n \leq nK'(t_0)\} &\leq e^{-n[t_0 K'(t_0) - K(t_0)]}, \quad t_0 \leq 0 \end{aligned}\right\} \quad \ldots(4)$$

Relation (4) gives Chernoff Bounds.

Illustration. Chernoff Bounds for $N(0, 1)$.

Here $\quad K(t) = \ln M(t) = t^2/2, \quad n = 1, \quad \lambda = K'(t) = t, \quad [K'(t_0) = 0 = \text{mean}]$

$\therefore \quad P\{X \geq \lambda\} \leq e^{-\lambda^2/2}, \; \lambda \geq 0; \quad P\{X \leq \lambda\} \leq e^{-\lambda^2/2}, \; \lambda \leq 0.$

10-40. *Cantelli Inequality*

Let X be a r.v. with $\mathsf{E}(X) = 0$ and Var $(X) = \sigma^2 < \infty$. Then

(a) $P\{X < t\} \geq t^2/(\sigma^2 + t^2), t > 0$. $\qquad$ (*b*) $P(X \leq t\} \leq \sigma^2/(\sigma^2 + t^2), \quad t < 0$.

Proof. (a) Since $0 = \mathsf{E}(X) = \int_{-\infty}^{\infty} x\,dF(x)$, it follows that

$$-t = \int_{-\infty}^{\infty} (x - t)\,dF(x) = \left[\int_{-\infty}^{t} + \int_{t}^{\infty}\right](x - t)\,dF(x) \geq \int_{-\infty}^{t} (x - t)\,dF(x), \quad [\text{2nd integral} \geq 0]$$

$$\therefore \quad t \leq \int_{-\infty}^{t} (t - x)\,dF(x) = \int_{-\infty}^{\infty} (t - x)\,I_t(x)\,dF(x). \quad [I_t(x) = 1 \text{ if } x \leq t, I_t(x) = 0, \text{ if } x > t]$$

or $\quad t \leq \mathsf{E}\,[(t - X)\,.\,I_t(X)]$

By squaring both sides and using Cauchy-Schwarz inequality $[\mathsf{E}\,(UV)]^2 \leq \mathsf{E}(U^2)\,\mathsf{E}(V^2)]$, we get

$$t^2 \leq [\mathsf{E}(t - X)\,.\,I_t(x)]^2 \leq \mathsf{E}(t - X)^2\,.\,\mathsf{E}[I_t^2(X)] = \mathsf{E}(t^2 - 2tX + X^2)\,.\,P(X < t)$$

$$= [t^2 - 2t\mathsf{E}(X) + \mathsf{E}(X^2)]\,.\,P(X < t) = (t^2 + \sigma^2)\,P(X < t). \qquad [\mathsf{E}(X) = 0, \;\; \mathsf{E}(X^2) = \sigma^2]$$

i.e. $\quad P(X < t) \geq t^2/(t^2 + \sigma^2)$. [Equivalently : $P(X \geq t) \leq \sigma^2/(\sigma^2 + t^2)$].

(*b*) When $t < 0$, then $-t > 0$, and for any variate Y, part (a) supplies

$$P\{Y \geq -t\} \leq \sigma_Y^2/(\sigma_Y^2 + t^2) \;\; \Rightarrow \;\; P\{X \leq t\} \leq \sigma_X^2/(\sigma_X^2 + t^2). \; [\text{Put } Y = -X, \sigma_Y^2 = \sigma_X^2]$$

General Case. If $\mathsf{E}(X) = \mu \neq 0$, then letting $Z = X - \mu$, $\;\mathsf{E}(Z) = 0$; $\sigma_Z^2 = \sigma_X^2$, we get

$$P\{X < t\} = P\{Z < t - \mu\} \geq (t - \mu)^2/[\sigma^2 + (t - \mu)^2], \qquad t > \mu.$$

$$P\{X \leq t\} = P\{Z \leq t - \mu\} \leq \sigma^2/[\sigma^2 + (t - \mu)^2], \qquad t < \mu.$$

10-41. Worked-out Problems

Example 1. If $\gamma_2 = (\mu_4/\sigma^4) - 3$ is known, show that it is possible to improve Chebyshev's inequality in the form

$$P\{|X - \mu| \geq k\sigma\} \leq \frac{\gamma_2 + 2}{\gamma_2 + 2 + (k^2 - 1)^2} = \frac{\mu_4 - \sigma^4}{\mu_4 - 2k^2\sigma^4 + k^4\sigma^4}.$$

where k is any positive real number greater than 1.

Solution. For any variate Y, we know by Cantelli inequality :

$$P\{Y \geq 1\} \leq \sigma_Y^2 / (1 + \sigma_Y^2) \qquad \ldots(1)$$

Let $Y = (Z^2 - 1)/(k^2 - 1), k > 1$, where $Z = (X - \mu)/\sigma$. Now $\mathrm{E}(Z^2) = 1, \mathrm{E}(Z^4) = \mu_4/\sigma^4$ and so

$$\text{Var}(Y) = \frac{\text{Var}(Z^2)}{(k^2-1)^2} = \frac{[\mathrm{E}(Z^4) - \mathrm{E}^2(Z^2)]}{(k^2-1)^2} \frac{(\mu_4/\sigma^4) - 1}{(k^2-1)^2} = \frac{\gamma_2 + 2}{(k^2-1)^2} \qquad \ldots(2)$$

From (1) and (2) we obtain

$$P\left\{\frac{Z^2 - 1}{k^2 - 1} \geq 1\right\} = P\{Z^2 \geq k^2\} = P\{|Z| \geq k\} \leq \frac{(\gamma_2 + 2)}{\gamma_2 + 2 + (k^2 - 1)^2} \qquad \ldots(3)$$

or $\qquad P\{|X - \mu| \geq k\sigma\} \leq (\gamma_2 + 2)/[\gamma_2 + 2 + (k^2 - 1)^2].$

Aliter. Let $p = P\{|X - \mu| \geq k\sigma\} = P\{Z^2 \geq k^2\} = P\{Z^2 - k^2 \geq 0\}$.

$$\therefore \quad p = P\left\{\frac{(Z^2 - k^2)(k^2 - 1)}{(k^2 - 1)^2 + (\gamma_2 + 2)} \geq 0\right\} \qquad (\because \gamma_2 + 2 = \beta_2 - 1 > 0)$$

$$= P\left\{\frac{(Z^2 - k^2)(k^2 - 1)}{(k^2 - 1)^2 + (\gamma_2 + 2)} + 1 \geq 0 + 1\right\} = P\left\{\frac{(k^2 - 1)(Z^2 - 1) + (\gamma + 2)}{(k^2 - 1)^2 + (\gamma_2 + 2)} \geq 1\right\};$$

$$p \leq \frac{(\gamma + 2)}{[(k^2 - 1)^2 + (\gamma_2 + 2)]} \qquad [\text{By §10-30 (1) using } \mathrm{E}(Z^2) = 1] \qquad \ldots(3)$$

Sharpness. The estimate admits improvement iff

$$\frac{(\gamma + 2)}{[(k^2 - 1)^2 + (\gamma_2 + 2)]} \leq \frac{1}{k^2} \text{ i.e. iff. } (k^2 - 1)(\gamma + 2) < (k^2 - 1)^2 \Rightarrow \beta_2 < k^2, \; (\because k^2 > 1)$$

Hence, whenever $\beta_2 < k^2$, the estimate (3) shall be improvement on Chebyshev's inequality.

Example 2. *Continuous Bimodal Distribution.* Let

$$f(x) = \frac{1}{2\sigma\sqrt{2\pi}}\left\{\exp\left[\frac{-(x-\mu)^2}{2\sigma^2}\right]+\exp\left[\frac{-(x+\mu)^2}{2\sigma^2}\right]\right\}, -\infty < x < \infty.$$

(i) Use Chebyshev's inequality to establish $P\{X^2 \geq t\} \leq (\mu^2+\sigma^2)/t.$

(ii) Obtain the stronger bounds : $P\{|X^2-\mu^2-\sigma^2| \geq t\} \leq (2\sigma^4+4\sigma^2\mu^2)/t^2.$

Solution. Firstly to find μ_4, we find m.g.f. of f.

$$M(t) = \frac{1}{2}(\sigma\sqrt{2\pi})^{-1}\int_{-\infty}^{\infty} e^{tX}\,[e^{-(x-\mu)^2/\sigma^2}+e^{-(x+\mu)^2/2\sigma^2}]\,dx$$

$$= \frac{1}{2}[\mathrm{E}(e^{tX})+\mathrm{E}(e^{tY})],\ [X \sim N(\mu,\sigma^2), Y \sim N(-\mu,\sigma^2)]$$

$$= \frac{1}{2}[e^{\mu t+(\sigma^2t^2/2}+e^{-\mu t+(\sigma^2t^2/2)}] = e^{\sigma^2t^2/2}\left[\frac{1}{2}(e^{\mu t}+e^{-\mu t})\right] = e^{\sigma^2t^2/2}\cosh(\mu t)$$

$$= \left(1+\frac{\sigma^2t^2}{2}+\frac{\sigma^4t^4}{2!\,4}+\dots\right)\left(1+\frac{\mu^2t^2}{2!}+\frac{\mu^4t^4}{4!}+\dots\right) \qquad \dots(1)$$

Now μ'_r = coeff. of $t^r/r\,!$ in (1). As $\mu'_1 = 0$, simple and central moments coincide.

$$\mu_2 = \mu^2+\sigma^2,\quad \mu_4 = \mu^4+6\sigma^2\mu^2+3\sigma^4.$$

(i) $P\{g(X) < t\} \leq \mathrm{E}[g(X)]/t.$ [§10-30 (1)]. Take $g(X) = X^2$, $\mathrm{E}(X^2) = \mu_2$, the result follows.

(ii) $P\{|(X-\mu')^2-\mu_2| \geq t\} \leq (\mu_4-\mu_2^2)/t^2$ [§10-32(3)]

Here $\mu_4-\mu_2^2 = 2\sigma^4+4\sigma^2\mu^2$, $\mu' = \mathrm{E}(X)$; the result follows.

Problems with Solutions Provided at the End of the Text

1*. If a positive variate X has mean μ and variance σ^2, show that

$$P\{X \geq t\} \leq \mathrm{E}[(X+c)^2]/(t+c)^2,\ t>0,\ \forall\ c>0.$$

If $\mu = 0$, deduce that $P\{X \geq t\} \leq \sigma^2/(\sigma^2+t^2)$.

2*. Let $X_1, X_2, \dots$ be a sequence of r.v.s and let $p_i \geq 0$, $i = 1, 2, \dots$ with $\Sigma p_i = 1$.

Prove that : $\sigma(\Sigma p_i X_i) \leq \Sigma p_i \sigma_i \leq \{\Sigma p_i \sigma_i^2\}^{1/2}$. $1 \leq i \leq \infty$, $[\sigma(X_i) = \sigma_i]$

3*. Let $X_1, \dots, X_n$ be n variates, $S_n = X_1+X_2+\dots+X_n$ and Var $(X_i) = \sigma_i^2$. Show that

$$\text{Var}\,(S_n/n) \leq \Sigma\sigma_i^2/n, \qquad (i = 1, 2, \dots, n).$$

4*. Let X be a r.v. whose all moments are finite. If m and n are positive integers

$$\mathrm{E}\{|X|^{(m+n)/2}\} \leq [\mathrm{E}(|X|^m).\mathrm{E}(|X|^n)]^{1/2} \qquad \dots(1)$$

Hence show that the curve $g(n) = \ell n\ \mathrm{E}(|X|^n)$ is convex in the plane of g and n.

10-50. *Definition of a Convex Function*

Let $\varphi : R \to R$. The function φ is called convex iff

$$\varphi(\lambda x + (1-\lambda)y) \le \lambda\varphi(x) + (1-\lambda)\varphi(y); \lambda > 0 \text{ and } x, y \in R.$$

Note. A sufficient condition for φ to be convex is that $\varphi''(x) \ge 0$. Geometrically, the graph $y = \varphi(x)$ lies on or above any of its tangents.

10-51. *Jensen's Inequality on Convex Function*

If φ is a convex function, X a random variable with finite mean $\mu = \mathsf{E}(X)$, then

$$\mathsf{E}[\varphi(X)] \ge \varphi[\mathsf{E}(X)].$$

Proof. Taylor's series for $\varphi(x)$ about $\mu = \mathsf{E}(X)$, with remainder after two terms is

$$\varphi(x) = \varphi(\mu) + (x-\mu)\varphi'(\mu) + \tfrac{1}{2}(x-\xi)^2 \varphi''(\xi)$$

where ξ is some value between x and μ. Since $\varphi''(\xi) \ge 0$ as φ is convex, we get $\varphi(x) \ge \varphi(\mu) + \varphi'(\mu)(x-\mu)$, true for all x. Hence, randomizing

$$\varphi(X) \ge \varphi(\mu) + (X-\mu)\varphi'(\mu).$$

Taking expectations and using $\mathsf{E}(X-\mu) \equiv 0$ yields

$$\mathsf{E}[\varphi(X)] \ge \varphi(\mu) + \varphi'(\mu)\mathsf{E}(X-\mu) = \varphi(\mu) \Rightarrow \mathsf{E}[\varphi(X) \ge \varphi[\mathsf{E}(X)].$$

Cor. If r.v. X has finite mean and $\psi(X)$ is a concave function then $\mathsf{E}[\psi(X)] \le \psi[\mathsf{E}(X)]$.

Proof. The result follows by observing that $-\psi$ is a convex function.

Illustrations :

1. $\varphi(x) = x^2$. $\varphi''(x) = 2 > 0$, hence φ is convex. It follows that

$$\mathsf{E}(X^2) \ge [\mathsf{E}(X)]^2 \Rightarrow \text{Var}(X) \ge 0.$$

2. $Q(x) = 1/x$ (Rect. Hyperbola), $\varphi''(x) = (2/x^3) > 0$, for $x > 0$. Hence

$$\mathsf{E}(1/X) \ge 1/\mathsf{E}(X), \text{ for all } X > 0.$$

3. $\varphi(x) = -\ln x$, $x > 0$. $\varphi''(x) = 1/x^2 < 0$, thus $(-\ln x)$ is convex function. Hence

$$\mathsf{E}[\ln(X)] \ge \ln[\mathsf{E}(X)], \quad X > 0$$

4. ***AM/GM Inequality*** : From Jenson's inequality, applied to $-\ln X$, we get

$$\mathsf{E}[\ln X] \le \ln[\mathsf{E}(X)] \Rightarrow \Sigma p_i \ln x_i \le \ln(\Sigma p_i x_i)$$

where $p_i = P(X = x_i)$, i = 1, 2, ..., n. Take $p_i = 1/n$ (equi-probable case)

$$\therefore \ln \Pi\, x_i^{p_i} \le \ln \Sigma p_i x_i \Rightarrow \Sigma p_i x_i \ge \Pi x_i^{p_i}$$

i.e. $(x_1 + \ldots + x_n)/n \ge (x_1 . x_2 \ldots x_n)^{1/n} \Rightarrow AM \ge GM.$

5.

$$\varphi(x) = e^{tx}, \varphi''(x) = t^2 e^{tx} \ge 0, \ \forall \ x. \text{ Hence}$$

$$\mathsf{E}(e^{tX}) \ge e^{t\mu} \Rightarrow M_X(t) \ge e^{\mu t}, \ [\mu = \mathsf{E}(X)].$$

10-52. *Lyapounov Inequality*

Let $E(|X|^b) < \infty$. Then for arbitrary a, b, $0 < a \le b$.

$$[E(|X|^a)]^{1/a} \le [E(|X|^b)]^{1/b}.$$

Proof. We note that, for $\lambda > 1$, $g(y) = (y)^\lambda$, $y \ge 0$ is convex. Hence by Jensen's Inequality :

$$E[(Y)^\lambda] \ge |E(Y)|^\lambda. \text{ Let } Y = |X|^a, \lambda = b/a \ge 1, \text{ then}$$

$$E(|X|^a)^{b/a} \le E(|X|^b) \Rightarrow [E|X|^a]^{1/a} \le [E|X|^b]^{1/b}.$$

Cor. If $X \ge 0$, then

$$E(X) \le [E(X^2)]^{1/2} \le [E(X^3)]^{1/3} \le [E(X^4)]^{1/4} \le \ldots \qquad \ldots(1)$$

Example : If X and Y are i.i.d. variates, with $E(X) = 0$ and $E(|X|^3) < \infty$, then $E(|X - Y|)^3 \le 4E(|X|)^3$.

Solution. $|X - Y|^3 = |X - Y|(X - Y)^2 \le (|X| + |Y|)(X^2 + Y^2 - 2XY)$
$= |X|^3 + |Y|^3 + X^2|Y| + Y^2|X| - 2XY.|Y| - 2XY.|X|$

Since $\quad E|X|^3 = E|Y|^3, E(XY.|X|) = E(Y).E(X.|X|) = 0$, etc.

$$\therefore \quad E(|X - Y|^3 \le 2E|X|^3 + 2E(X^2)E|Y| = 2E|X|^3 + 2E(X^2)E|X| \le 4E(|X|^3)$$

$$\left[\because E|X| \le E^{1/3}|X^3|,\ E(X^2) \le [E|X|^3]^{2/3}, \text{by Cor. to § 10-52}\right]$$

10-60. *Hölder's Inequality*

Let $p \ge 1$, $q \ge 1$, such that $p^{-1} + q^{-1} = 1$, and X, Y be positive variates. Then

$$E(XY) \le [E(X^p)]^{1/p}.[E(X^q)]^{1/q}. \qquad \ldots(1)$$

Proof. The function $Q(x) = \ell n\, x$, $x > 0$ is *concave*. That is, for each λ, $0 \le \lambda \le 1$,

$$\lambda\, \ell n\, x + (1 - \lambda)\, \ell n\, \lambda\, y \le \ell n\, [\lambda x + (1 - \lambda) y].$$

By antilogarithm : $x^\lambda . y^{1-\lambda} \le \lambda x + (1 - \lambda)\, y. \qquad \ldots$(i)

Put $\lambda = p^{-1}$, $1 - \lambda = q^{-1}$, $x = X^p / E(X^p)$, $y = Y^q / E(Y^q)$, then (i) yields

$$\frac{X}{[E(X^p)]^{1/p}} . \frac{Y}{[E(Y^q)]^{1/q}} \le \lambda \frac{X^p}{E(X^p)} + (1 - \lambda) \frac{Y^q}{E(Y^q)}. \qquad \ldots(ii)$$

We take expected value of both sides of (ii), clear the fractions and instantly obtain (1).

Cor. $[E(XY)]^2 \le E(X^2).E(Y^2)$. $(p = q = 2)$ [Cauchy-Schwarz Inequality].

10-61. *Minkowski's Inequality*

For every integer $k \ge 1$, and any r.v. X and Y with $E(X^k) < \infty$, $E(Y^k) < \infty$.[This means X and Y belong to L^k space : a generalized space]

$$[\mathsf{E}|X+Y|^k]^{1/k} \le (\mathsf{E}|X|^k)^{1/k} + (\mathsf{E}|Y|^k)^{1/k}. \qquad ...(1)$$

Proof. For $k = 1$, result (1) is trivially true. Now

$$|X+Y|^k = |X+Y|^{k-1}.|X+Y| \le |X+Y|^{k-1}.(|X|+|Y|)$$

$$\therefore \quad \mathsf{E}|X+Y|^k \le \mathsf{E}|X+Y|^{k-1}.|X| + \mathsf{E}(|X+Y|^{k-1}.|Y|)$$

$$< (\mathsf{E}|X|^{(k-1)s})^{1/s}.(\mathsf{E}|X|^k)^{1/k} + [\mathsf{E}|X+Y|^{(k-1)s}]^{1/s}.(\mathsf{E}\,|Y|^k)^{1/k}, \text{ [by Hölders inequality]} \quad ...(2)$$

where $k^{-1} + s^{-1} = 1$, and hence $(k-1)\,s = k$. Dividing both sides of (2) by $(\mathsf{E}|X+Y|^k)^{1/s}$ we obtain

$$[\mathsf{E}|X+Y|^k]^{(1-s^{-1})} \le (\mathsf{E}|X|^k)^{1/k} + (\mathsf{E}|Y|^k)^{1/k}. \quad [1 - s^{-1} = k^{-1}]$$

This is the result (1).

10-70. *Gurlands' Inequality*

Let φ and ψ be continuous monotonic functions of a r.v. X which are both non-decreasing (or both non-increasing) and for which $\varphi(X) \ge 0$, $\psi(X) \ge 0$. If expectations exist, then

$$\mathsf{E}[\varphi(X).\psi(X)] \ge \mathsf{E}[\varphi(X)].\mathsf{E}[\varphi(X)] \qquad ...(1)$$

If φ is non-decreasing and ψ non-increasing or vice versa, then

$$\mathsf{E}[\varphi(X).\psi(X)] \le \mathsf{E}[\varphi(X)].\mathsf{E}[\varphi(X)] \qquad ...(2)$$

Proof. Let φ and ψ be both non-decreasing, then for x and y in their domain of definitions, $y \ge x \Rightarrow \varphi(y) \ge \varphi(x)$ i.e. $\varphi(x) \ge 0$. Similarly, $y \ge x \Rightarrow \psi(y) - \varphi(x) \ge 0$; hence

$$[\varphi(y) - \varphi(x)]\,[\psi(y) - \psi(x)] \ge 0. \qquad ...(i)$$

If φ and ψ both decrease, then for $y \ge x$, we get $\varphi(y) \le \varphi(x)$, $\psi(y) \le \varphi(x)$ and these give $\varphi(y) - \varphi(x) \le 0$, $\psi(y) - \varphi(x) \le 0$, which once again provide (i).

Hence if $\varphi\uparrow$, $\psi\uparrow$ or $\varphi\downarrow$, $\psi\downarrow$, we always obtain for random variables X and Y

$$[\varphi(Y) - \varphi(X)]\,[\psi(Y) - \psi(X)] \ge 0. \qquad ...(A)$$

Let X and Y be i.i.d. variates then from (A), taking expectations

$$\mathsf{E}\{\varphi(Y)\,\psi(Y) + \varphi(X)\,\psi(X) - \varphi(X)\,\psi(Y) - \varphi(Y)\,\psi(X)\} \ge 0$$

or $$\mathsf{E}\{\varphi(Y)\,\psi(Y)] + \mathsf{E}[\varphi(X)\,\psi(Y)] - \mathsf{E}[\varphi(X)]\,\mathsf{E}[\psi(Y)] - \mathsf{E}[\varphi(Y)]\,\mathsf{E}[\psi(X) \ge 0] \qquad ...(B)$$

Further : $\mathsf{E}[\varphi(Y)\,\psi(Y)] = \mathsf{E}[\varphi(X)\,\psi(X)] \quad [\because X, Y \text{ are i.i.d}]$

Making substitutions into (B), cancelling factor 2, we get

$$\mathsf{E}[\varphi(X)\,\psi(X)] \ge \mathsf{E}[\varphi(X)]\,.\mathsf{E}[\psi(X)] \qquad ...(1)$$

This establishes the result (1). The result (2) follows by noting that if $\varphi\uparrow\ \psi\downarrow$ or $\varphi\downarrow$, $\psi\uparrow$, then instead of (A) we shall have

$$[\varphi(Y) - \varphi(X)]\,[\psi(Y) - \psi(X)] \le 0 \qquad ...(C)$$

which leads to (2), using the i.i.d nature of X, Y.

Extension. If $\varphi_1, \varphi_2, \ldots, \varphi_n$ are continuous monotone functions of a variate X which are all non-decreasing (non-increasing) and for which $\varphi_i(X) \geq 0$, $i = 1, 2, \ldots, n$, then

$$\mathsf{E}[\varphi_1(X).\varphi_2(X)\ldots\varphi_n(X)] \geq \mathsf{E}[\varphi_1(X)].\mathsf{E}[\varphi_2(X)]\ldots\mathsf{E}[\varphi_n(X)].$$

Some Illustrations

(i) $\mathsf{E}(X^{r-1}) \leq \mathsf{E}(X^r).\mathsf{E}(X^{-1}), r > 0$. Take $\varphi(X) = X^r$, $\psi(X) = X^{-1}$ in (2).

(ii) Take $\varphi(X) = \psi(X) = (X-\mu)^2$, then (1) provides $\mathsf{E}(X-\mu)^4 \geq [\mathsf{E}(X-\mu)^2]^2 \Rightarrow \beta_1 \geq 1$.

(iii) Take $\varphi(X) = \psi(X) = |X - \mu|$; then (1) provides

$$\mathsf{E}[(X-\mu)^2] \geq [\mathsf{E}(|X-\mu|)^2 \Rightarrow \sigma_X^2 \geq [\delta(\mu)]^2 \Rightarrow \text{S.D.} \not> \text{M.D.}$$

(iv) Take $\varphi(X) = X^\alpha, \psi(X) = X^\beta$ in (1) or (2) to get

$$\mathsf{E}(X^{\alpha+\beta}) \geq \mathsf{E}(X^\alpha)\mathsf{E}(X^\beta), \alpha, \beta > 0; \quad \mathsf{E}(X^{\alpha+\beta}) \leq \mathsf{E}(X^\alpha)\mathsf{E}(X^\beta) \ \alpha\beta < 0.$$

In particular, $\mathsf{E}(X^0) \leq \mathsf{E}(X)\mathsf{E}(1/X) \Rightarrow \mathsf{E}(X) \geq 1/\mathsf{E}(X^{-1})$.

$$\mathsf{E}(X^0) \leq \mathsf{E}(X^2)\mathsf{E}(X^{-2}) \Rightarrow \mathsf{E}(X^2) \geq 1/\mathsf{E}(X^{-2}).$$

In general : $\quad \mathsf{E}(X^{-n}) \leq 1/\mathsf{E}(X^n), n = 1, 2, 3, \ldots$

(v) Take $\varphi(X) = e^{t_1X}, \psi(X) = e^{t_2X}$, so that $\varphi(X)\,\psi(X) = e^{(t_1+t_2)X}$ Use (1) :

$$\mathsf{E}[e^{(t_1+t_2)X}] \geq \mathsf{E}(e^{t_1X})\mathsf{E}(e^{t_2X}) \Rightarrow M(t_1+t_2 : X) \geq M(t_1 : X)\,M(t_2 : X).$$

Exercise 10(b)

1. Let $\mathsf{E}(X) = 0$, and $\mathsf{E}(X^k) = \mu_k$, for a variate X. Using the relation $\mathsf{E}(AX^a + BX^b + CX^c)^2 \geq 0$, establish

$$\begin{vmatrix} \mu_{2a} & \mu_{a+b} & \mu_{a+c} \\ \mu_{b+a} & \mu_{2b} & \mu_{b+c} \\ \mu_{c+a} & \mu_{c+b} & \mu_{2c} \end{vmatrix} \geq 0.$$

Hence or otherwise show that Pearson's beta coefficients satisfy the relation $\beta_2 - \beta_1 - 1 \geq 0$.

Deduce also that $\beta_2 \geq 1$.

2. Show that $P\{|X| > a\} > [\mathsf{E}(X^2) - a^2]/b^2$, if $|X| < b$.
3. Let X be a positive variate, $\mathsf{E}(X) = \mu$ and Var $(X) = \sigma^2 < \infty$. If k is a number lying between 0 and 1, show that $P\{X > k\mu\} > (1-k)^2/\mathsf{E}(X^2)$.
4. The p.d.f. $f(x)$ of a variate X is symmetrical about $X = 0$, and has a single mode at $X = 0$, $-a \leq x \leq a$. Show that $\mu_{2r} < a^{2r}/(2r+1)$.
5. A r.v. X takes on the values 1, 2, 3, ... and $\mathsf{E}(X) < \infty$. Prove that : $\mathsf{E}(X) \geq 2 - p$, $\mathsf{E}(X) \geq 3 - 2p_1 - p_2, \ldots$ where $p_k = P(X = k)$. Generalize this inequality.
6. (a) A continuous variate $X \geq 0$ has a non-increasing p.d.f. f. Show that $f(x) \leq 2\mathsf{E}(X)/x^2$, for all $x > 0$.
   (b) Let $g(X) \geq 0 \ \forall\ x$ and $g(X) \geq k$ for $\forall\ x \in J =]a, b[$. Show that $P\{X \in J\} \leq k^{-1}\,\mathsf{E}[g(X)]$.
7. ***C_r-Inequality.*** If $\mathsf{E}(|X|^r) < \infty$, $\mathsf{E}(|Y|^r) < \infty$, then

$$\mathsf{E}(|X+Y|^r) \leq C_r[\mathsf{E}(|X|^r) + \mathsf{E}(|Y|^r)] \text{ where } C_r = 1, \text{ if } 0 < r \leq 1 \text{ and } C_r = 2^{r-1}, \text{ if } r > 1.$$

8. Let $g(x)$ be non-decreasing for $x > 0$ and $g(x) = g(-x) \geq 0$. Show that, for every $c > 0$, $P\{\,|X| \geq c\} \leq \mathsf{E}[g(X)]/g(c)$. Further, if g is bounded, i.e. $|g(x)| \leq K < \infty$, then $P\{\,|X| \geq c\} \geq \{\mathsf{E}[g(X)] - g(c)\}/K$.
Deduce : $P\{\,|X| \geq c\} \geq [\mathsf{E}g(X) - g(c)]/g(K)$, if $P\{\,|X| \leq K\} = 1$.

9. If m.g.f. $M(t)$ of a variate exists for every $t > 0$, show that

$$P\{tX > s^2 + \ell n\, M(t)\} < e^{-s^2} \quad \forall \quad s > 0.$$

10. Let $X_1, \ldots, X_n$ be completely independent variates, such that $\mathsf{E}(X_i) = \mu_i$ and $\mathrm{Var}\,(X_i) = \sigma_i^2 < \infty$, $|X_i - \mu_i| \leq B$, $1 \leq i \leq n$. Let further, $S_n = X_1 + X_2 + \ldots X_n$, $\mathsf{E}(S_n) = M$, $\mathrm{Var}\,(S_n) = \sigma^2$. Prove that $P\{\,|S_n - M| \geq \mu\,\sigma\} \leq 2\exp\{-\mu^2/2\,[1 + (\mu B/2\sigma)]^2$, where $\mu\ (> 0)$ is such that $\mu \leq \sigma/B$. Show further that, if $\mu_i = 0$ for all i, then $\ell n\mathsf{E}(e^{tX}) \leq \frac{1}{2}t^2\sigma^2\left(1 + \frac{1}{3}Bt\,e^{Bt}\right)$

11. For the positive r.v.s. X and Y, show that $\mathrm{Cov}\,(X, Y/X) \leq \mathrm{Var}\,(\sqrt{Y})$.

12. Prove that $\mathrm{Cov}\,(X, Y) \leq \mathrm{Var}\,[(X + Y)/2]$. If X and Y are positive variates, then this inequality generalizes to : $\mathrm{Cor}\,(X, Y) \leq \mathrm{Var}\,[X^k + Y^k)/2]^{1/k}, 0 \leq k \leq 1$.

13. For the non-degenerate independent variates X and Y show that
(i) $\mathsf{E}(X/Y)^r \geq \mathsf{E}(X^r)/\mathsf{E}(Y^2), r = 1, 2, \ldots$
(ii) $\mathrm{Var}\,(X/Y) \geq \mathrm{Var}\,(X)/\mathrm{Var}\,(Y)$ if $\mathsf{E}(X) = \mathsf{E}(Y) = 0$, and $\mathsf{E}(Y^{-1})$ exists.

14. If $\mathsf{E}(X) = 0$, $\mathrm{Var}\,(X) = \sigma^2$ and $|X| \leq m$, then $\mu'_r < v'_r \leq m_{r-2}\sigma^2, r \geq 2$.

15. Let X be a non-negative random variables and $\mathsf{E}(X_r) = M_r$. Show that, with usual notation : $M_r \geq (M_{r/s})^s \geq (M_1)^r + [M_{r/s} - (M_1)^{r/s}]^s, r \geq s \geq 1. M_r \geq \mu^r + \sigma^r$, when $r \geq 2$.

16. X and Y are independent continuous variates with c.d.f.s $F(x)$ and $G(x)$. Define : X is "stochastically larger" than Y if $F(t) \leq G(t), \forall\, t$. Prove that if this is $P\{X \geq Y\} \geq 1/2$.

17. (a) Let $f(x)$ be the p.d.f. of a continuous variate X satisfying $f^2(x) \geq f(x+1)\,f(x-1)$. Show that $g(r) = (\mu'_r/r\,!)^{1/r}$ is a decreasing function of r.
(*b*) Let $X \geq 0$ with $\mathsf{E}(X) = 100$. Comment on the bounds for the following :
(i) $\mathsf{E}(X^5)$, (ii) $\mathsf{E}(\sqrt{x})$, (iii) $\mathsf{E}(\ell n\,X)$, (iv) $e^{-tX}, t > 0$.

18. Let the p.d.f. of a variate $Y \in (a, b)$ be a decreasing function of Y. If $0 < a < b$, show that $(1 + r^{-1})^r a^r P(Y \geq a) \leq \mu'_r \leq (r+1)^{-1}\,\{a^r\,P(Y < a) + (b^{r+1} - a^{r+1})\,P(Y \geq a)/(b-a)\}$. Deduce Chebyshev-type inequalities.

19. If X is a unimodal variate, deduce the followings from Lemma to §10-22, p.360.

(i) $P\{|X - x_0| \geq k\tau\} \leq 4/9k^2;\ \forall\ k \geq 0.$ [Gauss' Inequality]

(ii) $P\{|X - \mu| \geq k\sigma\} \leq \{4(1+s^2)/9(k - |s|)^2\}, \quad \forall\ k > |s|.$

where x_0 is modal value, $s = (\mu - x_0)/\sigma$, [Pearson's measure of skewness] and $\tau^2 = \mathsf{E}(X - x_0)^2$.

[**Deductions.** $\tau^2 = \mathsf{E}(X - x_0)^2 = \mathsf{E}[(X-\mu) + (\mu - x_0)]^2 = \mathsf{E}(X-\mu)^2 + 2(\mu - x_0)\,\mathsf{E}(X-\mu) + (\mu - x_0)^2 = \sigma^2 + (\mu - x_0)^2$.

$$\therefore \quad \tau^2 = \sigma^2 + s^2\sigma^2 = (1+s^2)\,\sigma^2.$$

Let $f(x)$ be the density of the continuous r.v. X. Since x_0 is the only mode (maximum) of density f, it is necessarily decreasing for $x > x_0$. Hence setting $Y = |X - x_0| > 0$, $\lambda = k\tau$ in (1) yields

$$k^2\tau^2\,P\{|X - x_0| > k\tau\} \leq (4/9)\,\mathsf{E}\,(X - x_0)^2 \Rightarrow P\{|X - x_0| \geq k\tau\} \leq (4/9k^2). \quad \ldots\text{(i)}$$

(ii) From Bienayme' non-central inequality : $P\{|X-b|\ge c\}\le[E(X-b)^n]/c^n$, using $n=2$, we get

$$P\{|X-x_0|\ge c\}\le \tau^2/c^2=(1+s^2)\sigma^2/c^2;$$

$$P\{|X-\mu|\ge c\}\le\sigma^2/c^2\} \Rightarrow P\{|X-\mu|\ge c\}\le P\{|X-x_0|\ge c\} \quad ...(1)$$

$$[P\{|X-\mu|\ge k\sigma]\le 1/k^2, P\{|X-\mu|\ge k'\sigma\}\le 1/k'^2]$$

$$\Rightarrow P\{|X-\mu|\ge k\sigma\}\le P\{|X-\mu|\ge k'\sigma\} \quad ...(2)$$

where $k'=k-|s|$, so that $k'<k$ and $(1/k'^2)>(1/k^2)$. Thus from (1) and (2)

$$P\{|X-\mu|\ge k\sigma\}\le P\{|X-x_0|\ge k'\sigma\}. \quad ...(3)$$

Put $k\tau=k'\sigma$ i.e. $k=k'\sigma/\tau$ in inequality (i) to obtain

$$P\{|X-x_0|\ge k'\sigma\}\le\frac{4\tau'^2}{9\sigma^2k'^2}=\frac{4(1+s^2)\sigma^2}{9\sigma^2(k-|s|)^2}=\frac{4(1+s^2)}{9(k-|s|)^2}. \quad ...(4)$$

From (3) and (4) : $P\{|X-\mu|\ge k\sigma\}\le P\{|X-x_0|\ge k'\sigma\}\le\{4(1+s^2)/9(k-|s|^2\}$. ...(ii)

20. Show that the variance of the weighted means $\Sigma w_i X_i/\Sigma w_i$, $1\le i\le n$ of n independent variates X_i is a minimum when the weights are inversely proportional to the corresponding variables. Also show that this minimum variance is H/n, where H is the harmonic mean of the variances $\sigma_i^2, 1\le i\le n$.

21. *Berge's Inequality*. Let $Z_i=(X_i-\mu_i)/\sigma_i$, where $\mu_i=E(X_i)$ and $\sigma_i^2=\text{Var}(X_i)$, $i=1,2$; Corr $(X_1,X_2)=\rho$. Show that $P\{\max(|Z_1|,|Z_2|>c\}\le\left[1+\sqrt{1-\rho^2}\right]/c^2$.

22. (a) Show that if X is a r.v. such that $P(a\le X\le b)=1$, then $E(X)$ and Var (X) exist and $a\le E(X)\le b$ and Var $(X)\le(b-a)^2/4$. Prove further that $\mu/H\le(b+a)^2/4ab$, the equality being attainable iff half of the frequency lies at each extreme variate-value.

(b) Let X be a discrete variate, $P(X=x_i)=p_i$, $1\le i\le n$ and let $\Sigma p_i(x)^{(r)}=a$, $\min_i(x_i)=c$. Show that

(i) $c\nless E(X)\nless b$, (ii) Var $(X)\le(c-b)^2/4$, (iii) Var $(X)\le(c-a)/(a-b)$.

23. Let $X_1,X_2,...,X_n$ be independent variates : $S(X_1,...,X_n)$ and $T(X_1,...,X_n)$ being any two statistics which are non-decreasing functions of each component. Use Gurland's inequality and induction on n to show that $E(ST)\ge E(S)\,E(T)$.

24. Let $X_1,X_2,...,$ be i.i.d variates with $P(X_i>0)=1$ and Var $(\ln X_i)=\sigma^2$ and $Y_n=\Sigma_{i=1}^n \ln X_i$. Show that

$$P\{\exp n[E(\ln Y_i)-\varepsilon]<X_1.X_2...X_n<\exp n[E(\ln Y_i)+\varepsilon]\}\ge 1-(\sigma^2/n\varepsilon^2).$$

Deduce : $P\{X_1X_2...X_n<e^{n\varepsilon}[E(Y_i)]^n\}\ge 1-(\sigma^2/n\varepsilon^2)$.

Pay no attention to what critics say ... a statue has never been set up in honour of a critic. (J Sibelius)

Reality leaves a lot to the imagination. (John Lenon)

Convergence. Limiting M.G.F. Laws of Large Numbers

11

11-10. *Some Modes of Convergence in Probability Theory*

We recall that a sequence of real numbers $\langle x_n \rangle$ tends to limit x, if given $\varepsilon > 0$, we can find an integer n_0 such that

$$| x_n - x | < \varepsilon, \quad \forall \ n \geq n_0. \qquad ...(i)$$

Now consider the sequence of the random variables $\langle X_n \rangle : X_1, X_2, ..., X_n$. For each experimental outcome ω, we get a sequence of numbers

$$X_1(\omega), X_2(\omega), ..., X_n(\omega). \qquad ...(1)$$

Thus, (1) represents a family of sequences. This means we can no longer use (i) to define the convergence of the entire family.

We shall now define some modes of convergence for a sequence of random variables.

(A) ***Convergence in Probability or Stochastic Convergence***

Let r.v. X and sequence $\langle X_n \rangle$ of random variables have common domain S. Let P be the prob. measure defined on S and let $A_n = \{s : | X(s) - X_n(s) | > \varepsilon, s \in S\}$.

If $\lim_{n \to \infty} P(A_n) = \lim_{n \to \infty} P\{| X - X_n | > \varepsilon\} = 0, \forall \varepsilon > 0,$

then X_n is said to converge in probability (or *converge in measure*) to X and we write

$X_n \xrightarrow{p} X$.

Note that the definition says that as $n \to \infty$, the value of X_n will be arbitrarily close to the value X, except perhaps, at some x-values whose total probability approaches zero.

(B) ***Convergence in Distribution or in Law***

Give a sequence of random variables $\langle X_n \rangle$ having distribution functions $\{F_n\}$ respectively and a random variables X with distribution function F, then X_n is said to converge in distribution (or *Converge in Law*) to X, iff $\lim F_n(x) = F(x)$ as $n \to \infty$ at every point x, which is a continuity point of $F(x)$. In this case, we write $X_n \xrightarrow{d} X$ (limit variate) or $X_n \xrightarrow{L} X$.

Note that $X_n \xrightarrow{d} X$ at almost all x. The phrase : *almost all x* means that $F_n(x)$ and $F(x)$ differe significantly on at most a countable number of points on the real line.

(C) ***Strong convergence or convergence almost surely (A.S.)***

Let r.v. X and sequence $\langle X_n \rangle$ of random variables have common domain S.

$$B = \{s : \lim_{n \to \infty} X_n(s) = X(s), \ s \in S\}.$$

If $P(B) = P\{\lim X_n = X\} = 1$, then X_n is said to converge to X almost surely (certainly) a.s. or almost everywhere except on a *set of measure* zero. Very often, the definition is recorded in the form :

$$\lim_{n\to\infty} P\left[\bigcup_{k\geq n}\{|X_k - X| > \varepsilon\}\right] = \lim_{n\to\infty} P\left\{\sup_{k\geq n} |X_n - X| > \varepsilon\right\} = 0, \ \forall \ \varepsilon > 0.$$

The set $\bigcup\{|X_k - X| > \varepsilon\}, k \geq n,$ is the set of all values x where X_k differs from X by more than ε for any $k > n$.

(D) ***Mean Convergence***

Ler r.v. X and sequence $\langle X_n \rangle$ of random variables have common domain S. If

$$\lim_{n\to\infty} \mathsf{E}\{|X_n - X|^r\} = 0, \quad r \geq 1$$

then $\langle X_n \rangle$ is said to converge in the rth-order mean to X, and we write $X_n \xrightarrow{\text{r.m.}} X$. (convergence in the rth mean).

Note. In practice we take $r = 2$ (quadratic mean) and obtain convergence in the mean-square sense (m.s.). Thus, the sequence $\langle X_n \rangle$ tends to X in the m.s. sense (or limit in the mean : (L.I.M.) if $\mathsf{E}\{|X_n - X|^2\} \to 0$ for $n \to \infty$.

Definition. For $X_1, X_2, ..., X_n \in L^1$, $\langle X_n \rangle$ convergent X in L^1, if

$$\lim_{n\to\infty} \mathsf{E}[|X_n - X|] = 0, \text{ written } X_n \xrightarrow{L^1} X.$$

Summary

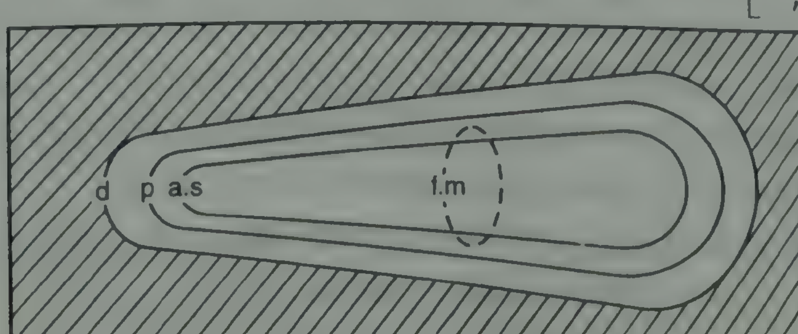
1. $X_n \xrightarrow{p} X$, if given $\varepsilon > 0$, $\lim_{n\to\infty} P\{|X_n - X| > \varepsilon\} = 0$.
2. $X_n \xrightarrow{d} X$, if $\lim_{n\to\infty} F_n(x) = F(x)$, $\forall\, x$ (continuity point of F).
3. $X_n \xrightarrow{\text{a.s.}} X$, if $P\left\{\lim_{n\to\infty} X_n = X\right\} = 1$.
4. $X_n \xrightarrow{\text{r.m.}} X$, if $\lim_{n\to\infty} \mathsf{E}(X_n - X)^r = 0, r \geq 1$.

11.11. *Relation between Various Modes of Convergence*

For every constant C and $(r > 1)$, the following results are true

1. $(X_n \xrightarrow{p} X) \Rightarrow (X_n \xrightarrow{d} X)$. **2.** $(X_n \xrightarrow{\text{a.s.}} X) \Rightarrow (X_n \xrightarrow{p} X)$ **3.** $(X_n \xrightarrow{\text{r.m}} X) \Rightarrow (X_n \xrightarrow{p} X)$.

4. $(X_n \xrightarrow{p} C) \Leftrightarrow (X_n \xrightarrow{d} C)$. **5.** $(X_n \xrightarrow{\text{r.m.}} C) \Rightarrow (X_n \xrightarrow{\text{a.s.}} C)$, $\left[\text{if } \sum_{n=1}^{\infty} \mathsf{E}(X_n - C)^r < \infty\right]$



Some implications of the modes of convergence.

Note. We can collect the implications :

(i) $(X_n \xrightarrow{\text{a.s.}} X) \Rightarrow (X_n \xrightarrow{p} X) \Rightarrow (X_n \xrightarrow{d} X)$.

(ii) $(X_n \xrightarrow{\text{r.m.}} X) \Rightarrow (X_n \xrightarrow{p} X) \Rightarrow (X_n \xrightarrow{d} X)$.

(iii) $(X_n \xrightarrow{\text{r.m.}} X) \Rightarrow (X_n \xrightarrow{\text{s.m.}} X)$, if $r > s \geq 1$.

Example : $(X_n \xrightarrow{\text{a.s.}} X) \Rightarrow (X_n \xrightarrow{p} X) \Rightarrow (X_n \xrightarrow{d} X)$.

Solution. (i) Write $Y_n = X_n - X$, so that $X_n \to X \;\Rightarrow Y_n = 0$. We further put

$A_n = \{|Y_n| < \varepsilon\}$, $T = \{\lim Y_n = 0, \text{ as } n \to \infty\}$.

$B_n = \{\text{Sup}\,|Y_k| < \varepsilon, k \geq n\}$; $B_{n+1} = \{\text{Sup}\,|Y_k| < \varepsilon,\ k \geq n+1\}$, so $B_n \subset B_{n+1}$, $B_n \subset A_n$

Because, $T \equiv \left\{\lim_{n\to\infty} Y_n = 0\right\} \subset \bigcup_{n=1}^{\infty} B_n$

$P\{T\} = 1$, by hypothesis (a.s.) and $\langle B_n \uparrow\rangle$, it follows that

$$1 = P\left\{\bigcup_{n=1}^{\infty} B_n\right\} = \lim_{n\to\infty} P(B_n), \text{ [Monotone } \uparrow \text{ Seq. of events]}$$

But $B_n \subset A_n \;\Rightarrow\; \lim P(B_n) \leq \lim P(A_n)$ which yields $\lim P(A_n) = 1$.

Thus a.s. cgc $\Rightarrow$ p–cgc. [cgc = convergence]

(ii) Let $\varepsilon > 0$ and $x \in \mathbb{R}$. Since $(X - X_n) \leq |X_n - X|$, so $X \leq X_n + |X_n - X|$. It follows that

$$\{X_n \leq x\} \Rightarrow \{X \leq x + \varepsilon\} \uplus |X_n - X| \geq \varepsilon\}.$$

$$\therefore \qquad P\{X_n \leq x\} \leq P\{X \leq x + \varepsilon\} + P\{|X_n - X| \geq \varepsilon\}.$$

Similarly, $\quad P\{X \leq x - \varepsilon\} \leq P\{X_n \leq x\} + P\{|X_n - X| \geq \varepsilon\}$.

Observe that $P\{|X_n - X| \geq \varepsilon\} \to 0$ as $n \to \infty$ (by hypothesis) and lim inf (..) $\leq$ lim sup f(..), so the two preceding equalities give, on letting $n \to \infty$, the conclusion

$$P\{X \leq x - \varepsilon\} \leq \liminf_{n\to\infty} P\{X_n \leq x\} \leq \lim_{n\to\infty} \text{Sup}\, P\{X_n \leq X\} \leq P\{X \leq x + \varepsilon\}.$$

For the distribution function $F_X(x) = P\{X \leq x\}$, continuous at x,

$$F_X(x - \varepsilon) = F_X(x - \varepsilon) = F_X(x) \text{ as } \varepsilon \to 0.$$

The continuous implications are established.

11-12. *Cauchy's Convergence Criterion (C.C.C.)*

Suppose $X_n \to X$ in some sense. In general, the limit variate X is unknown, and in order to test for convergence, we use Cauchy's Convergence Criterion :

$$|X_n - X_{n+m}| \to 0, \text{ for } n \to \infty \text{ and any } m > 0. \qquad \ldots(1)$$

If limit (1) exists in 'p' or 'a.e.' or in 'd' or in rth mean sense then X_n converges in the corresponding sense. For example, if given $\varepsilon > 0$, we can find n_0 such that

$$\mathsf{E}[|X_{m+n} - X_n|^2] < \varepsilon, \text{ for } n > n_0, \text{ and any } m > 0,$$

then $\langle X_n\rangle$ convergence in the mean square sense, etc.

11-20. *Illustrations for Various Modes of Convergence*

(A) Case of *convergence in probability*

1. Let X be any random variable and consider the sequence $\langle X_n \rangle$ of random variables such that $X_n = X + (1/n)$. Then $P\{| X_n - X | > \varepsilon\} = P\{| 1/n | > \varepsilon\}$.

Clearly, $\lim_{n\to\infty} P\{|X_n - X| > \varepsilon\} = 0,\ \forall\ \varepsilon > 0 \Rightarrow X_n \xrightarrow{p} X$.

2. Let X be a degenerate random variable : $P\{X = \mu\} = 1$. Let X_i $(i = 1, 2, ...)$ be i.i.d. variates with common mean μ and S.D. σ. Introduce Z_n by

$$Z_n = n^{-1}(X_1 + X_2 + ... + X_n),\ E(Z_n) = \mu,\ \text{Var}(Z_n) = \sigma^2 / n.$$

Then $P\{|Z_n - \mu| > \varepsilon\} \le \sigma^2 / n\varepsilon^2$. (Chebyshev's inequality)

Thus, $\lim_{n\to\infty} P\{|Z_n - \mu| > \varepsilon\} = 0$. But $\lim_{n\to\infty} P\{|Z_n - X| > \varepsilon\} = \lim_{n\to\infty} P|Z_n - \mu| > \varepsilon\} = 0$.

Hence, $Z_n \xrightarrow{p} X$, or for this degenerecy $Z_n \xrightarrow{p} \mu$.

3. Let us imagine the half open unit interval (0, 1] segmented into 10 subintervals $I_1 = (0, 0{\cdot}1]$, $I_2 = (0{\cdot}1, 0{\cdot}2]$, ..., $I_{10} = (0{\cdot}9, 1]$; or segmented into 100 subintervals : (0·01], (0·01 0·02], ..., (0·99, 1·0] or segmented into 1000 sub-intervals, or segmented into an infinite number of equal subintervals. Now let $X \sim U(0, 1)$ and consider $X_n = X + \delta(I_n)$, where

$$\delta(I_n) = 1 \text{ if } x \subseteq (I_n),\ \delta(I_n) = 0 \text{ if } x \notin I_n.$$

Clearly, $| X_n - X | = 0$, except for $x \in (I_n)$. Now, since X is $U(0, 1)$ and $P(x \in I_n)$ = length of I_n, hence

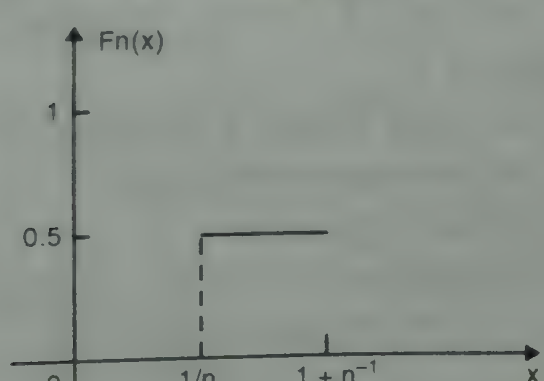
$$P\{| X_n - X | > \varepsilon\} = P\{x \in I_n\} = \text{length } I_n \Rightarrow \lim_{n\to\infty} P\{| X_n - X | > \varepsilon\} = \lim_{n\to\infty} I_n = 0 \Rightarrow X_n \xrightarrow{p} X.$$

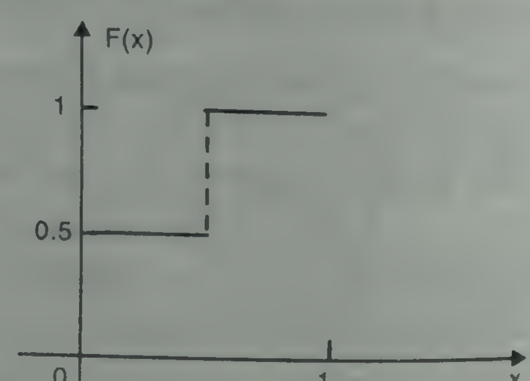
(B) Case of *Convergence in distribution*

1. Define a variate X by $P(X = 0) = 1/2 = P(X = 1)$. For any integer n, let variate X_n be defined by $X_n = 1 + (1/n) - X$. Since there is one-one correspondence between X_n and X, we find that the density for X_n is given by

$$P\{X_n = 1 + (1/n)\} = \tfrac{1}{2} = P\{X_n = (1/n)\}$$

The distribution function $F_n(x)$ and $F(x)$ are shown as under :





$X_n \xrightarrow{d} X$

As $n \to \infty$, we see that $F_n(x) \to F(x)$ at all points x except at the point $x = 0$ and $x = 1$. Thus $X_n \xrightarrow{d} X$.

2. Define variates X_n, $(n = 1, 2, \ldots)$ such that $F_n(x) = 1 - e^{-\lambda_n x}, x > 0,\ F_n(x) = 0,\ x \le 0$. We assume that $\lambda_n \to \lambda$ as $n \to \infty$. If r.v. X is such that $F(x) = 1 - e^{-\lambda x}, x > 0, F(x) = 0, x \le 0$, then we have $X_n \xrightarrow{d} X$.

(C) *Case of convergence almost surely*

1. Consider a sequence $\langle X_n \rangle$ of random variables with

$$P\{X_n = 1/n\} = 1/2 = P\{X_n = -1/n\}.$$

For $i > j$, $|X_i| > |X_j|$ so that $\{|X_i| > \varepsilon\} \subset \{|X_i| > \varepsilon\}$ whence the sequence $\langle |X_n| > \varepsilon \rangle$ is a decreasing sequence. It then follows that $\bigcup_{i=n}^{\infty}\{|X_i| > \varepsilon\} = \{|X_n| > \varepsilon\}$.

Choosing $n > 1/\varepsilon$, i.e. $\varepsilon > 1/n$, we observe that

$$P\left[\bigcup_{i=n}^{\infty}\{|X_i| > \varepsilon\}\right] = P\{|X_n| > \varepsilon\} \le P\{|X_n| > 1/n\} = 0. \Rightarrow \text{Thus, } X_n \to 0 \text{ almost surely.}$$

2. Let $\langle X_n \rangle$ be a sequence of independent variates such that

$$P\{X_n = 1\} = p_n, \quad P\{X_n = 0\} = 1 - p_n.$$

Let $0 < \varepsilon \le 1$, then $P\{|X_n| \ge \varepsilon\} = P\{X_n = 1\} = p_n$. Consider

$$P\left\{\bigcup_{j=n}^{\infty} |X_j| \ge \varepsilon\right\} \le \sum_{j=n}^{\infty} P\{|X_j| \ge \varepsilon\} = \sum_{j=n}^{\infty} p_j.$$

Thus, if $\sum p_j$ $(j = 1, 2, \ldots)$ is convergent, then $\lim_{n\to\infty} P\left\{\bigcup_{j=n}^{\infty} |X_j| \ge \varepsilon\right\} = 0$

i.e. $X_n \to 0$, almost surely, if $\sum p_n$ is convergent.

Remark. Let $r > 1$. Choosing $p_n = 1/n^r$, we may get several concrete situations.

(D) Case of *Convergence in the Mean*

1. Consider a sequence $\langle X_n \rangle$ of independent variates defined by

$$P\{X_n = 1\} = 1/n,\ P\{X_n = 0\} = 1 - (1/n),\quad n = 1, 2, \ldots$$

Here $E(|X_n - 0|^r) = E(|X_n|^r) = (1/n) \to 0$ as $n \to \infty$.

It follows that $X_n \to 0$ in rth mean $[X_n \xrightarrow{\text{r.m.}} 0]$.

2. Consider a sequence of variates $\langle X_n \rangle$ defined by

$$P(X_n = 0) = 1 - (1/n^2),\ P(X_n = n) = 1/n^2,\ n = 1, 2, 3, \ldots$$

Here $E\{|X_n - 0|^2\} = E(|X_n|^2) = 0.\ (1 - n^{-2}) + n^2(1/n^2) = 1$.

Thus, lime $\{|X_n - 0|^2\} \ne 0$, so $X_n \not\to 0$ in mean square.

***Example* 1.** Give an example of a sequence of r.v.s. which converges in distribution, (i) but not in probability, (ii) but not in any other mode of convergence.

Solution. (i) Let $P(Y = \pm 1) = 1/2$ and define the sequence $X_1, X_2, ..., X_n, ...$ by

$$X_n = Y \text{ if } n \text{ is odd; } \; X_n = -Y \text{ if } n \text{ is even.}$$

Obviously, $X_n \xrightarrow{d} Y$ as $n \to \infty$, because each variable in the sequence has the same distribution. However, $X_n - Y = 0$, if n is odd; $X_n - Y = -2Y$, if n is even.

$\therefore \; P\{| X_{2m} - Y | > 1\} = P\{| -2Y | > 1\} = P\{| Y | > 1/2\} = 1 \; \forall \; m.$

This proves that $X_n \xrightarrow{p} Y$.

(ii) Let $X \sim$ Ber (p), $p = 1/2$. Let $X_1, X_2, ...$ be identical variates given by $X_n = X, \forall \; n$.

The X's are certainly not independent, but $X_n \xrightarrow{d} X$. Define $Y = 1 - X$. Then $Y \sim$ Ber (p), $p = 1/2$. Obviously, $X_n \xrightarrow{d} Y$, because X and Y have the same distribution $[P(X = 1) = 1/2 = P(X = 0)]$. However, $X_n \xrightarrow{p} Y$ in any other mode of convergence because $| X_n - Y | = 1$, always.

***Example* 2.** If $X_1, X_2, ...$ is a sequence of r.v.s. and $X_n \xrightarrow{\text{r.m.}} X$, then $X_n \xrightarrow{p} X$ also. Is the converse true ?

Solution. (i) We have $P\{| Y | > c\} \le \mathsf{E}(Y^r)/c^r$, [Bienayme' Inequality]. Take $Y = X_n - X$ to obtain

$$P\{| X_n - X | > c\} \le \mathsf{E}[(X_n - X)^r]/c^2 \to 0 \text{ (by hypothesis). This proves } X_n \xrightarrow{p} X.$$

The converse to this result is untrue. Here is an example $X_n \xrightarrow{p} 0$ but $X_n \xrightarrow{\text{r.m.}} 0$

(ii) Let a sequence $\langle X_n \rangle$ have the p.m.f. : $P(X_n = 0) = 1 - n^{-1}$, $P(X_n = n) = n^{-1}$.

Then, given $\varepsilon > 0$ and for all large values of n,

$P\{| X_n | > \varepsilon\} = P\{X_n = n\} = 1/n \to 0$ as $n \to \infty$. Thus $X_n \xrightarrow{p} 0$.

$\mathsf{E}[(X_n - 0)^r] = \mathsf{E}(X_n^r) = 0 \,.\, (1 - n^{-1}) + n^r \,.\, n^{-1} = n^{r-1} \to \infty$ as $n \to \infty$. $(r \ge 2)$.

Thus $X_n \xrightarrow{\text{r.m.}} 0$.

Exercise. Verify the following results :

(a) If $X_n \xrightarrow{\text{a.s.}} X$ and $Y_n \xrightarrow{\text{a.s.}} Y$ then $X_n + Y_n \xrightarrow{\text{a.s.}} X + Y$.

(b) If $X_n \xrightarrow{p} X$ and $Y_n \xrightarrow{p} Y$, then $X_n + Y_n \xrightarrow{p} X + Y$.

(c) If $X_n \xrightarrow{r} X$ and $Y_n \xrightarrow{r} Y$, then $X_n + Y_n \xrightarrow{r} X + Y$.

(d) If $X_n \xrightarrow{d} X$ and $Y_n \xrightarrow{d} Y$, then $X_n + Y_n \xrightarrow{d} X + Y$.

(e) If $X_n \xrightarrow{d} X$ and $Y_n \xrightarrow{p} c$, then $X_n + Y_n \xrightarrow{d} X + c$.

(f) If $X_n \xrightarrow{L^1} X$ and $Y_n \xrightarrow{L^1} Y$, then $(X_n + Y_n) \xrightarrow{L^1} (X+Y)$.

(g) If $X_n \xrightarrow{d} X$ and $Y_n \xrightarrow{p} c$, then $X_n . Y_n \xrightarrow{d} cX$.

(h) If $X_n \xrightarrow{d} X$ and $Y_n \xrightarrow{p} c$, then $X_n / Y_n \xrightarrow{d} X/c$. [Cramer's theorem]

11-30. Continuity (or Limit) Theorem for M.G.F.

Let the variate X_n have the c.d.f. $F_n(x)$ and m.g.f. $M_n(t)$, that exists for $|t| < h$, for all n. If there exists a c.d.f. $F(x)$ with corresponding m.g.f. $M(t)$, defined for $|t| \le h_1 < h$, such that $\lim M_n(t) = M(t)$, as $n \to \infty$, then X_n has a limiting distribution with c.d.f. $F(x)$.

Continuity (or Limit) Theorem for Ch. Function

Let $X, X_1, X_2, \ldots$ be r.v.s. with ch. functions $\phi, \phi_1, \phi_2, \ldots$ Then $X_n \xrightarrow{d} X$ iff $\phi_n(t) \to \phi(t)\ \forall\ t \in R$.

***Example* 1.** If $X_n \xrightarrow{d} X$, then $aX_n + b \xrightarrow{d} aX + b$.

Solution.
$$\phi(t : aX_n + b) = e^{ibt}\,\phi(at : X_n) \to e^{ibt}\,\phi\,(at : X) \text{ as } n \to \infty.$$
$$\phi(t : aX : b) = e^{ibt}\,\phi\,(at : X) = \lim \phi\,(t : aX_n + b).$$

***Example* 2.** Let $X_1, X_2, \ldots,$ be indep. Cauchy r.v.s. Show that $\overline{X}_n = (S_n / n) \xrightarrow{d} X$, as $n \to \infty$.

Solution. $\phi(t : \overline{X}_n) = \phi(t : X_1)$. [See §8-71]. Hence $\phi(t : \overline{X}_n) \to \phi(t : \overline{X}_1)$ as $n \to \infty$. Thus $\overline{X} \xrightarrow{d} X_1$.

Polya's Theorem

If $F_n \to F$ and F is continuous, then the convergence is uniform, i.e.,
$$\lim_{n\to\infty} \sup |F_n(x) - F(x)| = 0.$$

Proofs of these three Theorems are beyond the scope of this book.

11-31. Worked-out Problems on Limiting m.g.f.

***Example* 1.** Find the m.g.f. of bin (n, p). Deduce that m.g.f. of standardized binomial is $(p.e^{qt/\sigma} + qe^{-pt/\sigma})^n$. Show that, as $n \to \infty$, this tends to $e^{t^2/2}$.

Solution. Let $X^* = (X - np)/\sqrt{npq} = (X - np)/\sigma$ where $\sigma = \sqrt{nq}$. Using mgf of X, we have
$$M(t : X^*) = M\left(t : \frac{X}{\sigma} - \frac{nq}{\sigma}\right) = e^{-npt/\sigma} M\left(\frac{t}{\sigma} : X\right) = e^{-npt/\sigma}(q + pe^{t/\sigma})^n = (pe^{qt/\sigma} + qe^{-pt/\sigma})^n.$$

For limit-process, we use exponential series :
$$pe^{qt/\sigma} + qe^{-pt/\sigma} = p\left\{1 + \frac{qt}{\sigma} + \frac{q^2t^2}{2!\sigma^2} + \frac{q^3t^3}{3!\sigma^3} + \ldots\right\} + q\left\{1 - \frac{pt}{\sigma} + \frac{p^2t^2}{2!\sigma^2} - \frac{p^3t^3}{3!\sigma^3} + \ldots\right\}$$
$$= 1 + (t^2/2n) + 0(1/n)$$

$\therefore \quad M(t : X^*) = [1 + (t^2/2n) + 0(1/n)]^n$.

$\lim\limits_{n\to\infty} M(t : X^*) = \lim\limits_{n\to\infty} [1 + (t^2/2n) + 0(1/n)] = e^{t^2/2}$, [m.g.f. of $N(0, 1)$], by Euler's limit.

It follows that, as $n \to \infty / X^*$ tends to be normally distributed.

***Example* 2.** Let $X_n \sim NB\ (k_n, q_n)$, $n = 1, 2, \ldots$ If $k_n \to \infty$, $p_n \to 0$, $(q_n = 1 - p_n)$ and $k_n p_n \to \lambda$ (fixed) as $n \to \infty$, find the limiting distribution of X_n.

Solution. $\quad M(t : X_n) = \{q_n / (1 - p_n e^t)\}^{k_n} = (1 - p_n)^{k_n} . (1 - p_n e^t)^{-k_n}$ [§ 8-16(6)]

$$= \left(1 - \frac{\lambda}{k_n}\right)^{k_n} . \left(1 - \frac{\lambda e^t}{k_n}\right)^{-k_n} \to e^{-\lambda} . e^{\lambda e^t} = e^{\lambda(e^t - 1)}, \text{ as } k_n \to \infty \quad \text{[By Euler's limit]}$$

Thus, $M(t : X_n) \to$ m.g.f. of Pois (λ), hence $X_n \xrightarrow{d} X$ where $X \sim$ Pois (λ).

Problems with Solutions Provided at the End of the Text

1*. Let $X \sim$ bin (n, p). If $n \to \infty$ and $p \to 0$ in such a way that $np = \lambda$ (constant), then $X \to$ Pois (λ).

2*. Let $X \sim$ Pois (λ). If $\lambda \to \infty$, then Poisson distribution tends to be Normal distributed.

3*. Let $X_n \sim$ gam (n, λ). Find the limiting distribution of $Y = (X_n/n)$ as $n \to \infty$.

4*. Let $X \sim NB^*\ (k, p)$. Show that, as $p \to 0$, $2pX \xrightarrow{d} Y$ where $Y \sim\ Y \sim \chi^2_{(2k)}$.

Exercise 11(a)

1. Let $P\{X_n = 1\} = 1/n$, $P\{X_n = 0\} = 1 - 1/n$. Find the limiting distribution of X_n.
2. Let $P\{X_n = \pm 1\} = 1/2$. Find the limiting distribution of $Y_n = \sum(X_k/2^k)$, $1 \le k \le n$, when X_i are indep. variates.
3. Find the limiting distribution of (X_n/n^2) when $X_n \sim \chi^2_{(n)}$..
4. Find the limiting distribution of $\overline{X}_n$ and X_i are standardized normal variates with $\rho_{ij} = \rho, i \ne j, i, j = 1, 2, \ldots, n$.
5. Let $X_n \sim$ gem (λ/n). Show that $(X_n/n) \sim$ Expo (λ) when $n \to \infty$.
6. Let $X_j \sim$ Pois (λ). Show that $(S_n - n\lambda)/\sqrt{n\lambda}$ has $N(0, 1)$ distribution as $n \to \infty$.
7. Let $X_n \sim\ N(\mu_n, \sigma_n^2)$. Show that $X_n \to N(\mu, \sigma^2)$, when $\mu_n \to \mu$ and $\sigma_n^2 \to \sigma^2$.

11-40. *Definition of the Weak Law of Large Numbers (W.L.L.N.)*

A sequence $\langle X_n \rangle$ of r.v.s. is said to satisfy the *weak law of large numbers* (W.L.L.N.) if

$$\lim_{n\to\infty} P\{|S_n / n - \mathrm{E}(S_n / n)| < \varepsilon\} = 1$$

for any $\varepsilon > 0$, where $S_n = X_1 + X_2 + \ldots + X_n$.

Note. This law is *weak* because it is formulated in terms of distributions alone, rather than convergence of r.v.s. themselves.

11-41. *Theorem : Weak Law of Large Numbers*

Let $\langle X_n \rangle$ be a sequence of random variables and let $S_n = X_1 + ... + X_n$, with $B_n = \text{Var}(S_n) < \infty$. Then, for any $\varepsilon > 0$,

$$\lim_{n\to\infty} P\left\{\left|\frac{S_n}{n} - \mathsf{E}\left(\frac{S_n}{n}\right)\right| < \varepsilon\right\} = \lim_{n\to\infty} P\{|\overline{X}_n - \mathsf{E}(\overline{X}_n)| < \varepsilon\} = 1$$

provided $(B_n/n^2) \to 0$, as $n \to \infty$.

Proof. Chebyshev's inequality applied to variable (S_n/n) gives

$$P\left\{\left|\frac{S_n}{n} - \mathsf{E}\left(\frac{S_n}{n}\right)\right| \geq \varepsilon\right\} \leq \frac{\text{Var}(S_n/n)}{\varepsilon^2} = \frac{B_n}{n^2\varepsilon^2}. \qquad \left[\text{Var}\left(\frac{S_n}{n}\right) = \frac{\text{Var } S_n}{n^2}\right]$$

As $n \to \infty$, the right side of this inequality tends to zero (by hypothesis). Since probability is non-negative, the above provides

$\lim P\{|(S_n/n) - \mathsf{E}(S_n/n)| \geq \varepsilon\} = 0$, or $\lim P\{|(S_n/n) - \mathsf{E}(S_n/n)| < \varepsilon\} = 1$, as $n \to \infty$ by complement rule.

Cor. Let $\overline{X}_n = S_n/n$ and $\mu = \mathsf{E}(S_n/n)$; then

$$\lim_{n\to\infty} P\{\overline{X}_n \leq k\} = \begin{cases} 0, & \text{if } k < \mu \\ 1, & \text{if } k > \mu \end{cases}$$

Proof. The statement of WLLN reads :

$$\lim_{n\to\infty} P\{|\overline{X}_n - \mu| < \varepsilon\} = 1, \text{ or } \lim_{n\to\infty} P\{|\overline{X}_n - \mu| \geq \varepsilon\} = 0. \qquad ...(i)$$

Since $\{\overline{X}_n \leq \mu - \varepsilon\} \subset \{|\overline{X}_n - \mu| \geq \varepsilon\}$, hence $P\{\overline{X}_n \leq \mu - \varepsilon \leq \mu - \varepsilon\} \leq P\{|\overline{X}_n - \mu| \geq \varepsilon\}$.
Proceeding to the limit and using (i), we get

$$\lim P\{\overline{X}_n \leq k\} = 0;\ k = \mu - \varepsilon,\ \varepsilon > 0 \Rightarrow k < \mu.$$

Further : $P\{\overline{X}_n \leq \mu + \varepsilon\} + P\{|\overline{X}_n - \mu| > \varepsilon\} \geq 1$, since region larger than sample space is covered. Now put $k = \mu + \varepsilon$, take limits as $n \to \infty$ and using (i) we get, $\lim P\{\overline{X}_n \leq k\} = 1$.

Remark. The condition $(B_n/n^2) \to 0$ as $n \to \infty$ is necessary as well as sufficient for WLLN to hold for a sequence of *uniformly* bounded variates. This fact is established in Theorem 11-61.

11-50. *Some Variations of the W.L.L.N.*

1. *Mean-square law of large numbers*

Let $X_1, X_2, ...$ be a sequence of indep. r.v.s., $\mathsf{E}(X_k) = \mu$, $\text{Var}(X_k) = \sigma^2, 1 \leq k < \infty$. Show that $(S_n/n) \xrightarrow{\text{r.m.}} \mu$.

Proof. As usual, $S_n = X_1 + X_2 + ... + X_n$, $\mathsf{E}(S_n) = \mathsf{E}(X_1 + ... + X_n) = n\mu$; $\mathsf{E}(S_n/n) = \mu$.

$$\text{Var}(S_n) = \text{Var}(X_1 + ... + X_n) = n\ \text{Var}(X_k) = n\sigma^2;\ \text{Var}(S_n/n) = \sigma^2/n.$$

$$\therefore \quad \mathsf{E}[(S_n/n) - \mu]^2 = \text{Var}(S_n/n) = \sigma^2/n \to 0 \text{ as } n \to \infty.$$

Thus $\quad (S_n/n) \xrightarrow{\text{r.m.}} \mu$ as $n \to \infty$. $(r = 2)$.

2(a). *Bernoulli's WLLN*

Let $\langle X_n \rangle$ be a sequence of Bernoulli trials with probability of success equal to p. If S_n is the number of successes in n trials, then

$$\lim_{n \to \infty} P\{|(S_n / n) - p| < \varepsilon\} = 1, \forall\, \varepsilon > 0. \qquad \ldots(1)$$

Proof. Since $X_i = 1$, with prob. p and $X_i = 0$, with prob. $q = 1 - p$, we have

$$E(X_i) = p, E(X_i^2) = p, \text{ Var } (X_i) = p(1-p) = pq.$$

$$\therefore \quad E(S_n) = E(X_1 + \ldots + X_n) = np, \text{ Var } (S_n) = npq \qquad [X_i \text{ are i.i.d. indicators}]$$

Thus $E(S_n / n) = p$, Var $(S_n / n) = pq / n$. By Chebyshev's inequality

$$P\left\{\left|\frac{S_n}{n} - p\right| \geq \varepsilon\right\} \leq \frac{pq/n}{\varepsilon^2} \leq \frac{1}{4n\varepsilon^2} \quad (\because \max pq = 1/4).$$

By Complement rule, we rewrite this

$$P\{|(S_n / n) - p| < \varepsilon\} \geq 1 - (1/4n\varepsilon^2).$$

As $n \to \infty$, this reduces to statement (1), as $p \not> 1$.

Remark. The probability of "the relative frequency being close to p" approaches 1 as the number of trials increases.

2(b). *Borel WLLN*

Let $R(A)$ be the relative frequency of an event A in n independent and identical trials of an experiment. Then

$$P\{|R(A) - P(A)| > \varepsilon\} \to 0, \text{ as } n \to \infty \left[P\left\{\left|\frac{X}{n} - p\right| > \varepsilon\right\} \to 0, \text{ as } n \to \infty\right] \qquad \ldots(1)$$

That is, for large n, $R(A)$ is likely to be close to $P(A)$.

Proof. Let X_j, $(i \leq j \leq n)$ be i.i.d. Bernoulli variates with parameter p, $0 < p < 1$.

Let $S_n = X_1 + \ldots + X_n$ and $\overline{X}_n = S_n/n$. Now $M(t : X_j) = E(e^{tX} j) = pe^t + qe^{0t} = q + pe^t$. So

$$M(t : \overline{X}_n) = M(t : S_n/n) = M(t/n : S_n) = [M(t/n : X_j)]^n = [q + p^{t/n}]^n = [1 + p(t/n) + 0(1/n)]^n$$

As $n \to \infty$, $M(t : \overline{X}_n \to e^{pt}$ (m.g.f. of degenerate r.v.). Hence as $n \to \infty$

$$P\{(S_n / n) \leq x\} \to F(x) = \begin{cases} 0, & x < p \\ 1, & x \geq p. \end{cases} \qquad \left[P\left\{\left|\frac{X}{n} - p\right| > \varepsilon\right\} \to 0, \text{ as } n \to \infty\right] \quad \ldots(2)$$

The result (2) is equivalent to Borel WLLN (1).

3. *Khintchine's WLLN*

Let $\langle X_n \rangle$ be a sequence of i.i.d. variates, with $E(X_i) = \mu < \infty$, $i = 1, 2, \ldots$ Then the WLLN holds, i.e. $P\{|(S_n / n) - \mu| > \varepsilon\} \to 0$, as $n \to \infty$.

Proof. Here $S_n = X_1 + X_2 + ... + X_n$, $\overline{X}_n = S_n / n$. Addition Theorem for Ch. functions give

$$\begin{aligned}\phi(t:\overline{X}_n) &= \phi(t:S_n/n) = \phi(t/n:S_n) = [\phi(t/n:X_i)]^n, \ (X_i : \text{i.i.d}) \\ &= [1 + i\mu(t/n) + 0(t/n)]^n \qquad \text{[By Taylor's Theorem]}\end{aligned}$$

As $n \to \infty$, the R.H.S. tends to Euler's limit : $e^{i\mu t}$, hence

$$\lim_{n\to\infty} \phi\,(t:\overline{X}_n) = e^{i\mu t}.$$

Now, exp $(i\mu t)$ is the Ch. function of a degenerate variate X such that $P(X = \mu) = 1$. The c.d.f. of X is given by $F(x) = 1$, if $X \ge \mu$, $F(x) = 0$, if $X < \mu$.

The c.d.f. is continuous everywhere except at $X = \mu$. Choose $\varepsilon > 0$, use Continuity Theorem to get

$$\lim_{n\to\infty} P\{\overline{X}_n \le \mu - \varepsilon\} = \lim_{n\to\infty} F_{\overline{X}_n}(\mu - \varepsilon) = F_X(\mu - \varepsilon) = 0. \qquad ...(i)$$

$$\lim_{n\to\infty} P\{\overline{X}_n \le \mu + \varepsilon\} = \lim_{n\to\infty} F_{\overline{X}_n}(\mu + \varepsilon) = F_X(\mu + \varepsilon) = 1. \qquad ...(ii)$$

$$\therefore \qquad \lim_{n\to\infty} P\{\overline{X}_n > \mu + \varepsilon\} = 0. \qquad \text{[by (ii)]} \qquad ...(iii)$$

From (i) and (iii) follows the result : $\lim\limits_{n\to\infty} P\{|\overline{X}_n - \mu| > \varepsilon\} = 0 = \lim\limits_{n\to\infty} P\{|(S_n/n) = \mu| > \varepsilon\}$.

4. *Bernstein's WLLN*

Let $\langle X_n \rangle$ be a sequence of variates for which Var $(X_n) = \sigma_n^2 < K$, for all i, where K is independent of n. If Cov $(X_i, X_j) = \sigma_{ij} \to 0$ as $|i - j| \to \infty$, (Asymptotic uncorrelatedness), then WLLN holds.

Proof. By Chebyshev's inequality : $P\{|\overline{X}_n - E(\overline{X}_n)| \ge \varepsilon\} \le [\text{Var}\,(\overline{X}_n)/\varepsilon^2], \forall\, \varepsilon > 0$ ...(1)

Further, $$\text{Var}\,(\overline{X}_n) = \frac{1}{n^2}\text{Cov}\left(\sum_{i=1}^{n} X_i, \sum_{j=1}^{n} X_j\right) = \frac{1}{n^2}\sum_i\sum_j \sigma_{ij}. \qquad ...(2)$$

As $\sigma_{ij} \to 0$ when $|i - j| \to \infty$, it follows that given δ, there is an integer m for which $|\sigma_{ij}| \le \frac{1}{2}\delta$, provided $|i - j| > m$. The covariance matrix $[\sigma_{ij}]$ contains n^2 elements, the set $A = \{|i - j| \le m\}$ has mn elements and hence the set $A' = \{|i - j| > m\}$ has $(n^2 - mn)$ elements. Thus, $|\sigma_{ij}| \le \frac{1}{2}\delta$, on A'. Further, by Chwarz's inequality, $|\sigma_{ij}| \le \sigma_i\sigma_j \le K$ for all i and j, in particular, $|\sigma_{ij}| \le K$ on A. Now, from (2).

$$\text{Var}\,(\overline{X}_n) = \frac{1}{n^2}\sum_i\sum_j |\sigma_{ij}| = \frac{1}{n^2}\left\{\sum_A |\sigma_{ij}| + \sum_{A'} |\sigma_{ij}|\right\} \le \frac{1}{n^2}\left\{\sum_A K + \sum_{A'} \frac{1}{2}\delta\right\}$$

i.e. $$\text{Var}\,(\overline{X}_n) \le \left\{mnK + \tfrac{1}{2}(n^2 - mn)\delta\right\}/n^2 \le \delta, \qquad ...(3)$$

provided $n \ge (2mk/\delta)$. Thus, as $n \to \infty$, Var $(\overline{X}_n) \to 0$, $(\delta \downarrow 0)$ and (1) provides

$$P\left\{|\overline{X}_n - E(\overline{X}_n)| \ge \varepsilon\right\} \to 0, \text{ as } n \to \infty, \text{ for all } \varepsilon > 0.$$

It follows that WLLN holds.

11-60. *Necessary and Sufficient Condition for WLLN*

Let $\langle X_N \rangle$ be any sequence of variates; write $S_n = X_1 + X_2 + \ldots + X_n$ and $Y_n = [S_n - \mathsf{E}(S_n)]/n$. A necessary and sufficient condition for the sequence $\langle X_n \rangle$ to satisfy the WLLN is that

$$\mathsf{E}[(Y_n^2)/(1+Y_n^2)] \to 0, \text{ as } n \to \infty. \qquad \ldots(1)$$

Proof. (I) Assume that (1) holds. Now for any positive numbers a and b, $a \geq b > 0$,

$$[a/(1+a)] \,.\, [(1+b)/b] \geq 1. \qquad \ldots(i)$$

Define the event $A = \{|\,Y_n\,| \geq \varepsilon\}$; then $\omega \in A \Rightarrow |Y_n|^2 \geq \varepsilon^2 > 0$. Replace a by Y_n^2, b by ε^2 in (i) to get an event :

$$B = \left\{\left(\frac{Y_n^2}{1+Y_n^2}\right)\left(\frac{1+\varepsilon^2}{\varepsilon^2}\right) \geq 1\right\}.$$

Since $\omega \in A \;\Rightarrow\; \omega \in B$, it follows $A \subseteq B$ whence $P(A) \leq P(B)$.

$$\therefore \quad P(A) \leq P\left\{\left(\frac{Y_n^2}{1+Y_n^2}\right) \geq \left(\frac{\varepsilon^2}{1+\varepsilon^2}\right)\right\} \leq \frac{\mathsf{E}[Y_n^2/(1+Y_n^2)]}{\varepsilon^2/(1+\varepsilon^2)} \qquad \text{[by Markov's Inequality]} \ldots(ii)$$

As $n \to \infty$, the right extreme member in (ii) tends to zero by assumptions (1). Hence

$$\lim_{n\to\infty} P\{|Y_n| \geq \varepsilon\} = \lim_{n\to\infty} P\left\{\left|\frac{S_n - \mathsf{E}(S_n)}{n}\right| \geq \varepsilon\right\} \to 0. \qquad \ldots(2)$$

This shows that the sequence $\langle X_n \rangle$ satisfies the WLLN.

(II) Now assume that (2) holds. Let Y_n be a continuous random variable with density $f(y)$. Write $A = \{|\,Y_n\,| \geq \varepsilon\}$, $A' = |\,Y_n\,| < \varepsilon\}$, then using $[y^2/(1+y^2)] < 1$,

$$\mathsf{E}\left\{\frac{Y_n^2}{1+Y_n^2}\right\} = \int_{-\infty}^{\infty} \frac{y^2 f(y)}{1+y^2}\,dy = \left(\int_A + \int_{A'}\right)\frac{y^2}{1+y^2} f(y)\,dy \leq \int_A f(y)\,dy + \int_{A'} y^2 f(y)\,dy \quad \left[\because \frac{y^2}{1+y^2} < y^2\right]$$

$$< P\{|\,Y_n\,| \geq \varepsilon\} + \varepsilon^2 P\{|\,Y_n\,| < \varepsilon\} \leq P\{|\,Y_n\,| \geq \varepsilon\} + \varepsilon^2. \quad (\because\ y^2 < \varepsilon^2 \text{ on } A') \ldots(iii)$$

As $n \to \infty$, the first number on the R.H.S. of (iii) tends to zero by assumption (2), and since $\varepsilon > 0$ is arbitrary small, we get $\lim \mathsf{E}[(Y_n^2)/(1+Y_n^2)] = 0$.

Cor. Let $\langle X_n \rangle$ be a sequence of independent variates, such that Var $(X_i) < \infty$, $i = 1, 2, \ldots$ and $(B_{n/n^2}) \to 0$ as $n \to \infty$, where B_n = Var $(S_n) = \Sigma$ Var (X_i). Then WLLN holds.

Proof. $\{Y_n^2/(1+Y_n^2)\} \leq Y_n^2 = (S_n'/n)^2$ where $S'_n = S_n - \mathsf{E}(S_n)$.

$$\therefore \quad \mathsf{E}[Y_n^2/(1+Y_n^2)] \leq (1/n^2)\,\text{Var}\,(S_n) = B_{n/n^2}, \qquad (X_i \text{ are indep.})$$

As $n \to \infty$, $(B_{n/n^2}) \to 0$, by hypothesis and so $\mathsf{E}[Y_n^2/(1+Y_n^2)] \to 0$ as $n \to \infty$. This in turn implies that WLLN holds.

Important Observation. The proof of Theorem 11-60 remains true even if $\mathsf{E}(X_k)$ does not exist. Define $Y_n = (S_{n/n})$ instead of $Y_n = [S_n - \mathsf{E}(S)]/n$. Example 1 illustrates this observation.

Example : Let X_1, X_2, ... be a sequence of i.i.d. Chy (1, 0) variates. Show that the WLLN does not hold for this sequence.

Solution. It is easy to show that $X = (S_n/n) \sim$ Chy (1, 0). [See Art. 9-71]. Now

$$\mathsf{E}\left(\frac{Y_n^2}{1+Y_n^2}\right) = \mathsf{E}\left(\frac{S_n^2}{n^2+S_n^2}\right) = \mathsf{E}\left(\frac{X^2}{1+X^2}\right) = \int_{-\infty}^{\infty} \frac{x^2}{1+x^2} \frac{1}{\pi} \frac{dx}{1+x^2}$$

$$= \frac{2}{\pi}\int_0^{\pi/2} \frac{x^2\,dx}{(1+x^2)^2} = \frac{2}{\pi}\int_0^{\pi/2} \sin^2\theta\, d\theta = \frac{1}{2}. \qquad [x = \tan\theta]$$

Since $\lim \mathsf{E}[S_n^2/(n^2+S_n^2)] \to 0$, WLLN does not hold for this sequence.

11-61. *WLLN for Bounded Variates*

If the variates X_i are *uniformly bounded,* then the necessary and sufficient condition for the WLLN to hold is that $\lim (B_n/n^2) = 0$, as $n \to \infty$, where $B_n = \text{Var } S_n = \text{Var } (X_1 + X_2 + ... + X_n)$.

Proof. Assume that the WLLN holds, then

$$\lim P\{|\overline{X}_n - \mathsf{E}(\overline{X}_n)| \geq \varepsilon\} = 0, \text{ as } n \to \infty. \qquad ...(1)$$

Since X_i are uniformly bounded, $|X_i - \mathsf{E}(X_i)| < C$, so $X_i - \mathsf{E}(X_i) < C$ and summing provides

$$S_n - \mathsf{E}(S_n) < nc, \text{ or } \overline{X}_n - \mathsf{E}(\overline{X}_n) < C \text{ or } W_n < C,$$

where $W_n = \overline{X}_n - \mathsf{E}(\overline{X}_n)$. Also define the events $A = \{|W_n| \geq \varepsilon\}$ and $A' = \{|W_n| < \varepsilon]$.

Further, $\quad \text{Var }(\overline{X}_n) = \mathsf{E}[\overline{X}_n - \mathsf{E}(\overline{X}_n)]^2 = \mathsf{E}(W_n^2);\ B_n = \text{Var }(n\overline{X}_n) = n^2 \text{ Var }(\overline{X}_n). \qquad ...(2)$

Let $F(w)$ be the c.d.f. of W_n and write W for W_n for neatness; then

$$\mathsf{E}(W^2) = \int_{-\infty}^{\infty} w^2 dF(w) = \int_A w^2 dF(w) + \int_{A'} w^2\, dF(w).$$

Now on A', $w^2 < \varepsilon^2$ and on A, $w^2 < C^2$ (boundedness), hence the above gives

$$\mathsf{E}(W^2) \leq C^2\int_A dF(w) + \varepsilon^2\int_{A'} dF(w) = C^2\, P(A) + \varepsilon^2 P(A'). \qquad ...(3)$$

From (2) we get, $(B_n/n^2) = \text{Var }(\overline{X}_n) = \mathsf{E}(W^2)$, and thus (3) becomes

$$(B_n/n^2) \leq C^2 P(A) + \varepsilon^2 P(A').$$

As $n \to \infty$, $P(A) \to 0$ by (1) and since $P(A') < 1$, we finally obtain :

$$\lim (B_n/n^2) \leq \varepsilon^2 \Rightarrow \lim (B_n/n^2) \to 0, (\varepsilon \downarrow 0) \text{ as } n \to \infty.$$

The converse fact : $\lim (B_n/n^2) \to 0$ implies holding of WLLN, has already been establ hed

11-62. Worked-out Problems

***Example* 1.** Show that the indep. sequence $P\{X_K = \pm 2^k\} = 1/2$ does not obey the WLLN.

Solution. Here $E(X_k) = 2^k\left(\frac{1}{2}\right) + (-2^k)\left(\frac{1}{2}\right) = 0$. $\sigma_k^2 = \text{Var}(X_k) = E(X_k^2) = 2^{2k}\left(\frac{1}{2}\right) + 2^{2k}\left(\frac{1}{2}\right) = 2^{2k} = 4^k$.

$$\text{Var}(S_n) = \sum_{k=1}^{n} (\sigma_k)^2 = \sum_{k=1}^{n} 4^k = \frac{4}{3}(4^n - 1).$$

$$\lim_{n\to\infty}\left(\frac{\text{Var}(S_n)}{n^2}\right) = \frac{4}{3}\lim_{n\to\infty}\left(\frac{4^n}{n^2}\right) = \frac{4}{3}\lim_{n\to\infty}\left(\frac{4^n \ln 4}{2n}\right) = \frac{2}{3}(\ln 4)^2 \lim_{n\to\infty} 4^n \to \infty.$$

where we have used L'Hospital Rule for evaluation of (∞/∞) form. Since the condition $\text{Var}(S_n)/n^2 \to 0$ is only sufficient, not necessary, it does not follows that the WLLN does not hold. Further testing is necessary. We follow m.g.f. technique.

$$M(t : X_k) = E(e^{tX_k}) = \frac{1}{2}e^{2^k t} + \frac{1}{2}e^{-2^k t}. \qquad \text{...(i)}$$

$$M(t : S_n/n) = M(t/n : S_n) = M(\theta, X_1 + \ldots + X_n),\ [\theta = t/n]$$

$$= M(\theta : X_1)\ldots M(\theta : X_n) = \Pi\{(e^{2^k\theta} + e^{-2^k\theta})/2\}, \qquad \text{[by (i)]}$$

$$= (1/2)^n \Pi(\rho^{2^k} + \rho^{-2^k}),\ \text{where } \rho = e^{\theta}$$

$$(2^n)M(t : S_n/n) = \Pi\{(\rho^{2^k} + \rho^{-2^k})(\rho^{2^k} - \rho^{-2^k})/(\rho^{2^k} - \rho^{-2^k})\}$$

$$= \Pi\{(\rho^{2^{k+1}} - \rho^{-2^{k+1}})/(\rho^{2^k} - \rho^{-2^k}),\ [1 \le k \le n]$$

$$= \frac{\rho^{2^2} - \rho^{-2^2}}{\rho^2 - \rho^{-2}} \cdot \frac{\rho^{2^3} - \rho^{-2^3}}{\rho^{2^2} - \rho^{-2^2}} \cdots \frac{\rho^{2^{n+1}} - \rho^{-2^{n+1}}}{\rho^n - \rho^{-2^n}}$$

$$= \frac{\rho^{(2)^{n+1}} - \rho^{-(2)^{n+1}}}{\rho^2 - \rho^{-2}} = \frac{e^{(2)_\theta^{n+1}} - e^{-(2)_\theta^{n+1}}}{e^{2\theta} - e^{-2\theta}} = \frac{\sinh (2)^{n+1}\theta}{\sinh(2\theta)}$$

So
$$M(t : S_n/n) = \frac{1}{2^n} \cdot \frac{\sinh[(2)^{n+1} . t/n]}{\sinh(2t/n)} \qquad \left[\text{Use } \sinh x = x + \frac{x^3}{3!} + \frac{x^5}{5!} \ldots\right]$$

$$= \frac{1}{2^n} \cdot \frac{\alpha + (\alpha^3/3!) + (\alpha^5/5!) + \ldots}{\beta + (\beta^3/3!) + (\beta^5/5!) + \ldots}$$

where $\alpha = 2^{n+1}\dfrac{t}{n}$, $\beta = \dfrac{2t}{n}$, $\left(\dfrac{\alpha}{2^n\beta}\right) = 1$.

So $M(t : S_n/n) = \left[1 + \left(\dfrac{\alpha^2}{3!}\right) + \left(\dfrac{\alpha^4}{5!}\right) + \ldots\right] \Big/ \left[1 + \left(\dfrac{\beta^2}{3!}\right) + \left(\dfrac{\beta^4}{5!}\right) + \ldots\right]$

As $n \to \infty$, $\beta \to 0$ and $\alpha \to \infty$. Thus $M(t : S_n/n) \to \infty$ as $n \to \infty$

By Continuity Theorem, we conclude that the WLLN does not hold.

Remark. We shall not repeat above expanded arguments in the following solutions. The reader shall supply the details.

***Example* 2.** If X_i can have only two values i^α and $-i^\alpha$, with equal probabilities, show that WLLN can be applied to $\{X_n\}$, if $\alpha < \frac{1}{2}$. Also discuss the case when $\alpha = 1$.

Solution. Here $E(X_i) = \frac{1}{2}i^\alpha + \frac{1}{2}(-i^\alpha) = 0$. Var $(X_i) = E(X_i^2) = \frac{1}{2}i^{2\alpha} + \frac{1}{2}i^{2\alpha} = (i)^{2\alpha}$

$$\frac{B_n}{n^2} = \frac{1}{n}\left(\frac{1^{2\alpha}}{n} + \frac{2^{2\alpha}}{n} - \ldots + \frac{n^{2\alpha}}{n}\right) < \frac{1}{n}\left(\frac{1^{2\alpha}}{1} + \frac{2^{2\alpha}}{2} + \ldots + \frac{n^{2\alpha}}{n}\right)$$

By Cauchy's First Theorem on limits,

$$\lim_{n\to\infty}\left[\left(\frac{1^{2\alpha}}{1} + \frac{2^{2\alpha}}{2} + \ldots + \frac{n^{2\alpha}}{n}\right)\Big/ n\right] = \lim_{n\to\infty}\left(\frac{n^{2\alpha}}{n}\right) = 0, \text{ if } 2\alpha < 1.$$

Hence $\lim (Bn/n^2) = 0$, as $n \to \infty$

It follows that under this condition $(\alpha < \frac{1}{2})$, WLLN holds.

Case $\alpha = 1$. Here $\sigma_i^2 = i^2$, so that Var $(S_n) = \sum i^2 = n(n+1)(2n+1)/6$. Thus $(B_{n/n^2}) \to \infty$.

Further, $E(|X_i|)^{1+\delta} = (i)^{1+\delta}, \delta > 0$ is unbounded. Hence $\langle X_i \rangle$ does not admit of WLLN.

***Example* 3.** Prove that the WLLN is applicable to the arithmetic mean of a sequence of independent variates X_k specified by $P\{X_k = \pm\sqrt{\ell n\, k}\} = \frac{1}{2}$

Solution. Here $E(X_k) = \frac{1}{2}(\ell n\, k)^{(1/2)} - \frac{1}{2}(\ell n\, k)^{1/2} = 0$. Var $(X_k) = E(X_k^2) = \frac{1}{2}(\ell n\, k) + \frac{1}{2}(\ell n\, k) = \ell n\, k$.

$$\therefore \qquad B_n = \Sigma\sigma_i^2 = \ell n 1 + \ell n 2 + \ldots + \ell n\, n = \ell n\, n!$$

$$\frac{B_n}{n^2} = \frac{1}{n}\left[\frac{\ell n 1}{n} + \frac{\ell n 2}{n} + \ldots + \frac{\ell n\, n}{n}\right] < \frac{1}{n}\left[\frac{\ell n 1}{1} + \frac{\ell n 2}{2} + \ldots + \frac{\ell n\, n}{n}\right]$$

Since $\lim(\ell n\, n/n) = 0$ as $n \to \infty$, by Cauchy's First Theorem in Limits, $(B_n/n^2) \to 0$ as $n \to \infty$.

Hence $\lim (B_n/n^2) = 0$, as $n \to \infty$.

This proves the applicability of the Weak Law for large numbers.

Problems with Solutions Provided at the End of the Text

1*. State WLLN. Let $X_1, X_2, \ldots, X_n$ be i.i.d. variates with mean μ, and variance σ^2 and as $n \to \infty$, $(X_1^2 + X_2^2 + \ldots + X_n^2)/n \xrightarrow{p} c$ for some constant c, $(0 \le c < \infty)$. Find c.

2*. If the i.i.d. variates $X_k (k = 1, 2, 3, \ldots)$ assume the value $2^{r - 2\log r}, (r = 1, 2, 3, \ldots)$ with probability $1/2^r$, examine if the WLLN holds for the sequence $\langle X_n \rangle$

3*. A sequence of independent variates $\langle X_{nr} \rangle$ is defined by

$$P\left\{X_n = \frac{(r+1)^{k/2}}{r}\right\} = \frac{r}{(1+r)^k}, k = 1, 2, 3, \ldots \qquad [r \text{ is a positive integer}]$$

Show that $\langle X_{nr} \rangle$ admits of WLLN.

4*. Let $X_1, X_2, \ldots$ be a sequence of i.i.d. $U(0, 1)$ variates. For the geometric mean $G_n = (X_1, X_2, \ldots, X_n)^{1/n}$, show that $G_n \xrightarrow{P} c$, where c is some constant. Find c.

5*. Sequences $\langle X_n \rangle$ and $\langle Y_n \rangle$ are defined by the laws : $X_j \sim U(0, j)$ and $Y_j \sim U(-j, j)$, $1 \le j \le n$. Do they obey WLLN ?

6*. Show that the following sequence of indep. variates does not obey WLLN :

$$P\{X_K = \pm(2k-1)^{1/2}\} = 1/2. \qquad (X_k : \text{independent})$$

7*. A variate X_k has the distribution

$$P(X_k = 0) = 1 - (2/3^{2k+2}),\ P(X_k = 3^k) = P(X_k = -3^k) = 3^{-(2k+2)}.$$

Does the WLLN hold for the independent variates sequence $\langle X_k \rangle$?

8*. Let $\langle X_n \rangle$ be a sequence of mutually independent variates such that

$$P\{X_n = \pm 1\} = \tfrac{1}{2}(1 - 2^{-n}),\ P\{X_n = \pm 2^{-n}\} = 2^{-n-1}.$$

Does the WLLN hold for this sequence ?

9*. (i) The sequence $\langle X_k \rangle$ is such that $\mathsf{E}(X_k^2) < A^2$, for all $k = 1, 2, \ldots$ and each variate X_k depends on variates with adjacent numbers. Show that the sequence obeys WLLN. (ii) Using Chebyshev's inequality, find the upper bound of

$$P\left\{\left|\tfrac{1}{2}(X_k + X_{k+1}) - \tfrac{1}{2}\mathsf{E}(X_k + X_{k+1}\right| \ge \sigma_k + \sigma_{k+1}\right\}.$$

10*. The variates $X_1, X_2, \ldots, X_n, \ldots$ have equal expectations and finite variations. Is the law of large numbers applicable to this sequence if all the covariance σ_{ij} are negative ?

11*. Let X_i be i.i.d. variates $i = 1, 2 \ldots$ with mean μ and variance $\sigma^2 < \infty$, and let $S_n = X_1 + \ldots + X_n$. Show that the WLLN does not hold for the sequence $\langle S_n \rangle$, but it holds for the sequence $\langle a_n S_N \rangle$ provided $na_n \to 0$

Exercise 11(b)

0. State the limit theorem for characteristic functions.

1. Examine whether the WLLN holds good for the following sequences of independent random variables :

(a) $P\{X_k = (-1)^k k\} = 6/\pi^2 k^2,\ k = 1, 2, 3, \ldots$

(b) $P\{X_k = \pm k^{-1/2}\} = \frac{1}{2}$.

(c) $P\{X_k = \pm k^{1/4}] = \frac{1}{2}$.

(d) $P[X_k = 1] = p_k, P\{X_k = 0\} = 1 - p_k$.

(e) $P\{X_i = k\} = 1/k^2\zeta(3), k = 1, 2, \ldots$ and $\zeta(3) = \Sigma(1/k^3)$.

(f) $P\{X_k = \pm k\} = \frac{1}{2}$. [No]

(g) $P\{X_k = \pm 1\} = k^{-1}, P\{X_k = 0\} = 1 - 2k^{-1}$.

(h) $P\{X_k = \pm\sqrt{3}) = 1/6, P\{X_k = 0\} = 2/3$.

(i) $P\{X_k = k^{-1/2}\} = 2/3, P\{X_k = -k^{-1/2}\} = 1/3.$

(j) $P\{X_k = \pm 2^k\} = 2^{-(2k+1)}, P\{X_k = 0\} = 1 - 2^{-2k}$. [Yes]

(k) $P[X_k = k^{-1}] = p_k, P\{X_k = 1 + k^{-1}\} = 1 - p_k.$

(l) $P\{X_k = \pm k\} = (1/2\sqrt{k}), P(X_k = 0) = 1 - (1/\sqrt{k})$. [No]

(m) $P\{X_k = \pm 2^k\} = 1/2^{k+1}, P(X_k = \pm 1) = \frac{1}{2}[1 - (1/2)^k]$.

2. Let $\langle X_k \rangle$ be a sequence of independent variates such that :

(a) $P\{X_k = \pm k^\alpha\} = \frac{1}{4}, P\{X_k = 0\} = \frac{1}{2}$. Determine α for which WLLN holds. [$\alpha < 1/2$]

(b) $P\{X_k = \pm 2k^\alpha\} = \frac{1}{4}$. Determine α for which WLLN holds. [$\alpha < 1/2$]

3. The following sequences are composed of pairwise independent variates. Show that WLLN holds for such sequences :

(a) $x_{ni} = \pm\sqrt{n}, 0$ (b) $x_{ni} = \pm\sqrt{\ell n\, n}$ (c) $x_{ni} = \pm na, 0$ (d) $x_{ni} = \pm na = 0$

$P(X_n = x_{ni}) = 1/n, 1 - 2/n$ $P(x_n = x_{ni}) = 1/2$ $P(X = x_{ni}) = 1/2n^2, 1 - 1/n^2$ $P(X = x_{ni}) = \frac{1}{2}n, 1 - \frac{1}{2}^{n-1}$

4. Verify by direct calculation that the WLLN holds for a sequence of independent $N(\mu, \sigma^2)$ variates.

5. Variates $X_1, X_2, \ldots$ are such that $E(X_k) = 0, P\{|X_k| < 4\} = 1, k = 1, 2, \ldots,$ and Var $(S_n / n) \to \frac{1}{2}$ as $n \to \infty$. Prove that X's cannot be independent and that WLLN does not hold for X's.

6. Let $X_1, X_2, \ldots$ be i.i.d. variates with p.m.f. $f(x) = \alpha^x/(1 - \alpha), x = 0, 1, 2, \ldots, 0 < \alpha < 1$. Show that WLLN holds for X's.

7. Let $X_1, X_2, \ldots$ be i.i.d. variates with c.d.f.

$F(x) = (1/2a^2)\,u(x + a) + (1 - 1/a^2)u(x) + (1/2a^2)\,u(x - a), [a > 2]$ where $u(x) = 1$, if $x \geq 0$; $u(x) = 0$, if $x < 0$ (unit step function). Does the WLLN hold for the X_i ?

8. Variates $X_1, X_2, \ldots$ are independent with $P\{X_n = \pm a_n\} = p_n$, $P\{X_n = 0\} = 1 - 2p_n$. Show that the sequence $\langle X_n \rangle$ obeys the WLLN if $n^{-2}(\Sigma p_i a^{i2}) \to 0$ as $n \to \infty$ $(i = 1, 2, \ldots, n)$.

9. Let $\langle X_n \rangle$ be a sequence of random variables such that X_k is independent of X_j, $j \neq k + 1, j \neq k - 1$. If Var $(X_k) < c$, (const.) for all k, show that WLLN holds for $\langle X_n \rangle$.

10. The p.d.f. of a variate Y is $f(y) = \frac{1}{2}, -1 \leq y \leq 1$. Show that the sequence $\langle X_n \rangle$ satisfies WLLN where $X_k = \sin(k\pi Y), k = 1, 2, \ldots$.

11. Let X_i, $i \leq i \leq n$ be uncorrelated variates with common mean μ and variances σ_i^2 respectively. If there exist constants $\alpha > 0$ and $\beta < 1$ such that $\sigma_i^2 < \alpha\beta^i$, then WLLN holds for the sequence $\langle X_n \rangle$.

12. In a sequence of independent games it is possible to win or lose a unit stake at the kth game with probabilities p_k and $1 - p_k$ respectively. If S_n denotes the total gain in the first n games, prove that $\lim_{n\to\infty} P[n^{-1} | S_n - E(S_n) | \geq \varepsilon\} = 0, \forall\, \varepsilon > 0$, irrespective of the values of the p_k. Hence strengthen the WLLN to show that for any P_k.

$$\lim_{n\to\infty} P\{|S_n - E(S_n)| \geq \varepsilon n^{1/2}\, \ell n\; n\} = 0.$$

13. Let $X_0, X_1, \dots, X_{n-1}$ be uncorrelated random variables with mean zero and common variance σ^2. Show that the WLLN holds for the sequence $\langle Y_k \rangle$ where

$$Y_k = \sum_{j=1}^{k} 2^{-j} X_{k-j}, \quad k = 1, 2, \dots, n.$$

14. Prove that the WLLN holds for a sequence of random variables in which each variate can depend only on variates with adjacent numbers, and all the variates contained in the sequence have finite variances and expectations.

15. A pack of cards numbered 1, 2, ..., n is shuffled and the cards are dealt one at a time. The variate X_i possesses the values 1 or 0 according as the ith card dealt has number i on it or not, and each card is equiprobable to appear at the ith place. Show that the WLLN holds for the sequence $\langle X_n \rangle$.

16. A sequence $\langle Z_n \rangle$ is uniformly bounded, i.e.

$$P\{a \le Z_n \le b\} = 1, n = 1, 2, \dots \ (a, b \text{ finite constants}).$$

Suppose that for any $\varepsilon > 0$, $\lim P\{|Z_n - m| \ge \varepsilon\} = 0$, as $n \to \infty$, $m \in [a, b]$. Show that $\lim \mathrm{E}(Z_n) = m$, as $n \to \infty$. Give an example to show that this result does not necessarily hold if the condition of uniform boundedness is removed.

17. Let $\langle X_n \rangle$ be a sequence of variates with $\sigma_n^2 \le A, \forall n$ and $\sigma_{ij} \to 0$ as $|i - j| \to \infty$. Show that the WLLN holds for this sequence.

18. *Bernstein polynomials.* Let $f : [0, 1] \to \mathbb{R}$ be a continuous map. Using WLLN, show that

$$\lim_{n\to\infty} \sum_{r=0}^{n} f\left(\frac{r}{n}\right)\binom{n}{r} x^n (1-x)^r \xrightarrow{\text{uniformly}} f(x), \ 0 \le x \le 1.$$

11-70. *Definition : Strong Law of Large Numbers (SLLN)*

A sequence $\{X_n\}$ of variates is said to satisfy the strong law of large numbers (SLLN) if

$$\{[S_n / n - \mathrm{E}(S_n / n)]\} \xrightarrow{\text{a.s.}} 0 \text{ as } n \to \infty.$$

Note. This law concerns intrinsically the convergence of random variables themselves rather than their distributions, hence it is a strong law.

1. ***Kolmogorov's SLLN***

Let $X_1, X_2, \dots$ be independent variates with $\mathrm{E}(X_i) = \mu_i$, $\mathrm{Var}(X_i) = \sigma_i^2 < \infty$. If $\Sigma[\sigma_k^2 / k^2] < \infty$, $k = 1, 2, \dots, \infty$, then SLLN holds for the sequence $\{X_n\}$.

2. ***Necessary and Sufficient Condition***

The existence of the expectation is a necessary and sufficient condition for the SLLN to hold for an i.i.d. sequence $\langle X_n \rangle$.

3. ***Borel's SLLN***

For a sequence of Bernoulli trials with constant probability p of success, the SLLN holds.

4. ***Difference between WLLN and SLLN***

We have (i) WLLN : If $\overline{X}_n \xrightarrow{p} \mu$; (ii) SLLN : If $\overline{X}_n \xrightarrow{\text{a.s.}} \mu$.

The WLLN states that for any fixed large value of n, say m, (S_m/m) is likely to be near μ. However, it does not say that (S_n/n) is bound to stay near μ $\forall\ n > m$. It means that the large values of $|(S_n/n) - \mu|$ can turn up infinitely often, may be at in frequent intervals. The SLLN shows that it cannot happen. In particular, it shows almost certainly that : $|(S_n/n) - \mu| > \varepsilon$ only a finite number of times.

Proof of one simple version of SLLN is now offered. Proofs of other results are beyond the scope of the present text.

11-71. Theorem : Strong LLN for Independent and Identical Variates

Let $X_1, X_2, ..., X_n, ...$ be a sequence of i.i.d. variates; each having a finite mean $\mu = E(X_j)$. Then

$$P\left\{\lim_{n\to\infty} \frac{X_1 + ... + X_n}{n} = \mu\right\} = 1.$$

Proof. We start with assuming that $\mu = 0$ and write $S_n = X_1 + X_2 + ... + X_n$. Then $S_n = \Sigma X_i = \Sigma X_j = \Sigma X_k = \Sigma X_r$ for i, j, k, r are dummies. Now

$$E(S_n^4) = E\{(\Sigma X_i)(\Sigma X_j)(\Sigma X_k)(\Sigma X_r)\} \qquad ...(1)$$

The expression above in R.H.S. consists of the terms of the types

$$X_i^4,\ X_i^3 X_j,\ X_i^2 X_j^2,\ X_i^2 X_i X_j,\ X_i X_j X_k X_r \qquad (i \neq j \neq k \neq r)$$

As $E(X_i) = 0$ and $X_i, X_j, ...$ are independent, we have

$$E(X_i^3 X_j) = E(X_i^3)E(X_j) = 0,\ \ E(X_i^2 X_i X_j) = E(X_i^2)E(X_i)E(X_j) = 0,\ \ E(X_i X_j X_k X_r) = 0.$$

Now, for a given pair i, j there shall be $\binom{4}{2} = 6$ terms in (1). Thus we can rewrite (1) as

$$E(S_n^4) = nE(X_i^4) + 6\binom{n}{2}E(X_i^2 X_j^2)$$

$$= n\lambda + 3n(n-1)E(X_i^2)E(X_i^2). \qquad [\lambda = E(X_i^4)]$$

From $\quad \mathrm{Var}(X_i^2) = E(X_i^4) - [E(X_i^2)]^2 \geq 0$ we get $[E(X_i^2)]^2 \leq \lambda$

$\therefore \quad E(S_n^4) \leq n\lambda + 3n(n-1)\lambda$

or $\quad E(S_n^4/n^4) \leq (\lambda/n^3) + (3\lambda/n^2)$

Extending $n \to \infty$, this yields

$$E\left\{\sum_{n=1}^{\infty} \frac{S_n^4}{n^4}\right\} = \sum_{n=1}^{\infty} E\left(\frac{S_n^4}{n^4}\right) < \infty. \qquad \text{(Convergent Series)}$$

This result implies that, with probability 1, $\sum_{n=1}^{\infty}\left(\frac{S_n^4}{n^4}\right) < \infty,$

For otherwise, the sum being infinite yields infinite expected value. As for convergent series, the nth term tends to zero as $n \to \infty$, hence, with probability 1,

$$\lim_{n\to\infty} \frac{S_n^4}{n^4} = 0, \text{ i.e. } \lim_{n\to\infty}\left(\frac{S_n}{n}\right)^4 = 0 \Rightarrow \lim_{n\to\infty}\left(\frac{S_n}{n}\right) = 0$$

Now if $\mu \neq 0$, we let $Y_j = X_j - \mu$ and apply the above arguments to sequence $\langle Y_j \rangle$ so that, with probability 1,

$$\lim_{n\to\infty} \sum_{j=1}^{n}\left(\frac{X_j - \mu}{n}\right) = 0, \Rightarrow \lim_{n\to\infty} \sum_{j=1}^{n} \frac{X_j}{n} = \mu, \text{ i.e., } \lim_{n\to\infty}\left(\frac{S_n}{n}\right) = \mu.$$

Exercise 11(c)

1. If $\mathrm{E}(X) < \infty$, then $\mathrm{E}(\overline{X}_n) = \mathrm{E}(X) = \mu$, and $P\left\{\lim_{n\to\infty} (\overline{X}_n) = \mu\right\} = 1.$
2. If $\mathrm{E}(\hat{S}_n^2) < \infty$, then $\mathrm{E}(\hat{S}_n^2) = \sigma^2$ and $P\{\lim_{n\to\infty} \hat{S}_n^2 = \sigma^2\} = 1.$
3. If $\mathrm{Var}\,(X_i) \le c < \infty, 1 \le i \le n$ and all X_i are independent, then $\langle X_n \rangle$ obeys SLLN.
4. Show that the SLLN remains valid when $\mathrm{E}(X_i) = \infty$.
5. Let A and B be two events in a probability space $(\Omega, \mathcal{F}, P)$. A random experiment is performed independently n times. Let T_n be the number of times the outcome is in A. Among these T_n outcomes in A, there are S_n outcomes that are also in B. Show that

$$P\left\{\lim_{n\to\infty} (S_n / T_n) = P(B / A)\right\} = 1, P(A) \neq 0.$$

Doubt is the father of invention.

I know nothing except the fact of my ignorance. (Socrates)

Bernoulli Binomial Distribution 12

12-10. *Bernoulli Trials*

A sequence of trials with the following properties is called a *sequence of Bernoulli's trials* :

1. The result of each trial can be classified into one of two categories, say success and failure (call them S and F).
2. $P(S) = p$: a constant for each trial $[P(F) = q = 1 - p]$.
3. Outcomes of each trial are independent of all the others.
4. The series of trials is performed a *fixed* number of times, say n.

12-11. *Bernoulli Binomial Distribution*

Let a series of n Bernoulli trials be performed, with $P(S) = p$, $P(F) = q$, $(q + p = 1)$. Let us determine the probability, $f(x)$, that exactly x successes occur in these n trials, $0 \le x \le n$. Let S_i be the event of occurring a success at the ith trial ($i = 1, 2, \ldots, n$). Then one sequence of events consisting of x successes and hence $(n - x)$ failure is

$$E_x = S_{i_1} . S_{i_2} \ldots S_{i_x} F_{j_1} \ldots F_{j_{n-x}}$$

Since each event is independent and $P(S_i) = p$, $P(F_j) = q$ for all i, j, we get

$$P(E_x) = P(F_{i_1}) P(F_{i_2}) \ldots P(F_{i_x}) P(F_{j_1}) \ldots P(F_{j_{n-x}}) = p^x q^{n-x}.$$

There are $[n! / (n - x) ! \ x !] = {}^nC_x$ mutually exclusive sequence of events for which exactly x successes occur. Since the order is irrelevent and only the number of successes is of interest, the probability $f(x)$ that exactly x successes will occur in a sequence of n Bernoulli trials is

$$f(x) = \binom{n}{x} q^{n-x} p^x, \quad x = 0, 1, 2, \ldots n.$$

The set of ordered pairs $\{x, f(x) \ ; \ 0 \le x \le n\}$ is the general binomial distribution. We shall often write $b\ (x \ ; n, p)$ for $\binom{n}{x} q^{n-x} p^x$, or more generally X is bin (n, p) or even X is $b\ (n, p)$.

Definition. A r.v. X is binomially distributed with parameters n and p if its probability law is

$$\boxed{f(x) = \binom{n}{x} q^{n-x} p^x, \quad 0 \le p \le 1, \ q = 1 - p, x = 0, 1, 2, \ldots, n.}$$

Note. n is called the *index* of B.D. (Binomial Distribution).

12-12. Simple Numericals on Probability Formula

***Example* 1.** The jury members decide independently and that each makes a correct decision with probab. p, $0 < p < 1$. If the decision of the majority is final, what is preferable : a 3 member jury or a single juror ?

Solution. Let X denote the number of jurors who decide correctly among a 3-member jury. Then X is bin $(3, p)$. For majority decision.

$$P(X \geq 2) = P(X=2)+P(X=3)$$

$$= \binom{3}{2} q p^2 + \binom{3}{3} p^3 = 3p^2 q + p^3 = 3p^2 - 2p^3.$$

Single juror decides correctly with probab. p, hence 3-member jury is preferable if $P(X \geq 2) > p$.
Thus $3p^2 - 2p^3 > p \Rightarrow 2p^2 - 3p + 1 < 0 \Rightarrow (p-1)(p-1/2) < 0$.

Thus, if $p > 1/2$, three-member jury is preferable. If $p < 1/2$, single juror is preferable. For $p = 1/2$, $P(X \geq 2) = 1/2$, so it makes no difference what option is implemented.

***Example* 2.** Suppose two-and four-engine planes fly if at least half of their engines work. If q is the probability of failure for a single engine and engines perform independently, find the values of q for which two-engine plane is to be preferred to the four-engine plane.

Solution. Let X be the number of engines that do not fail and let S_k denote successful flight with k-engine plane. Then

$$P(S_2) = P(X \geq 1) = 1 - P(X=0) = 1 - \binom{2}{0} q^2 = 1 - q^2.$$

$$P(S_4) = P(X \geq 2) = 1 - P(X=0) - P(X=1) = 1 - \binom{4}{0} q^4 - \binom{4}{1} q^3 p = 1 - 4q^3 + 3q^4, \ (p = 1-q).$$

$$P(S_2) \geq P(S_4) \Rightarrow 1 - q^2 \geq 1 - 4q^3 + 3q^4, \text{ or } q^2(1 - 4q + 3q^2) \leq 0, \text{ i.e. } q^2(1-q)(1-3q) \leq 0.$$

If $q = 0$, $q = 1$, $q = 1/3$, the above is a strict equality and the two kinds of planes have equal chances of successful flights.

If $1/3 < q < 1$, 2-engine plane is preferred. However if $0 < q < 1/3$, then the 4-engine plane is to be preferred.

***Example* 3.** Teams A and B are playing a series of games, each of which is independently won by team A with probability p and by team B with probab. $q = 1 - p$. The winner of the series is the first team to win 4 games. Find $E(X)$, where X is the number of games played.

Solution. Since series is won by 4 successes, there cannot be more than 7 games between A and B. We define the events for A and similarly for B.

$$A_k^n = \{\text{Team } A \text{ wins } k \text{ games out of } n \text{ games played}\}$$
$$A_k = \{\text{Team } A \text{ wins } k\text{th game}\}$$

$$\therefore \ P(X=4) = P\{A_4^4 \uplus B_4^4\} = P(A_4^4) + P(B_4^4) = p^4 + q^4$$

$$P(X=5) = P\{A_3^4 A_5 \uplus B_3^4 B_5\} = P(A_3^4 \,.\, A_5) + P(B_3^4 \,.\, B_5)$$

$$= \left\{\binom{4}{3} p^3 q\right\} p + \left\{\binom{4}{3} q^3 p\right\} q = 4\, pq\,(p^3 + q^3).$$

$$P(X=6) = P\{A_3^5 \,.\, A_6 \uplus B_3^5 B_6\} = P(A_3^5 A_6) + P(B_3^5 B_6)$$

$$= \left\{\binom{5}{3} q^2 p^3\right\} p + \left\{\binom{5}{3} p^2 q^3\right\} q = 10\, p^2 q^2\,(p^2 + q^2).$$

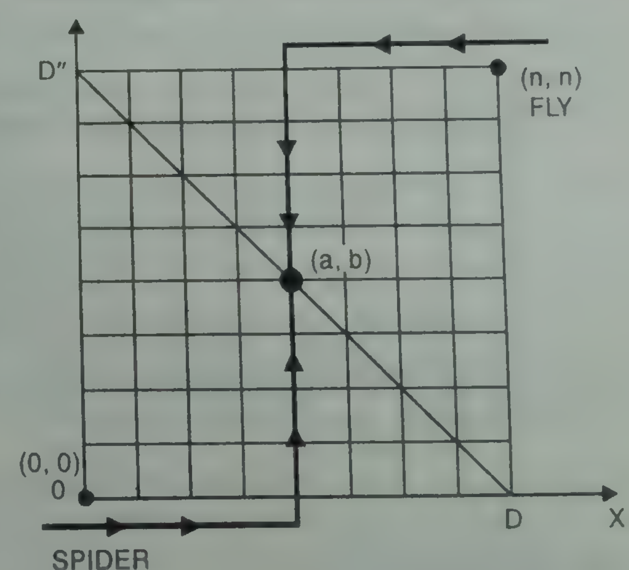
$$P(X=7) = P\{A_3^6 \,.\, A_7 \uplus B_3^6 B_7\} = P(A_3^6 A_7) + P(B_3^6 B_7)$$

$$= \left\{\binom{6}{3} q^3 p^3\right\} p + \left\{\binom{6}{3} p^3 q^3\right\} q = 20\, p^3 q^3, \quad (p+q=1).$$

$$\mathsf{E}(X) = \Sigma k\, P(X=k) = 4(p^4+q^4) + 5[4pq(p^3+q^3)] + 6\,[10\,p^2 q^2 (p^2+q^2)] + 7\,(20\,p^3 q^3).$$

Note. If $p = q = 1/2$, then $\mathsf{E}(X) = 8p^4 + 40p^5 + 260p^6 = 93/16 = 5.813$.

***Example* 4.** A spider and a fly are situated at the corners $(0, 0)$ and (n, n) of a rectangular grid. The spider walks only north or east, the fly only south or west. They take their steps simultaneously to an adjacent vertex of the grid. Show that, if the successive steps are independent and equally likely to go in each of the two possible directions, the probability that they will meet is $\binom{2n}{n}\left(\frac{1}{2}\right)^{2n}$.



Solution. Suppose the insects meet after k steps. If the spider walks a steps east and b steps north, it is at (a, b) with $a + b = k$. The fly must walk $n - a$ steps west and $n - b$ steps south, it is at (a, b) with $(n - a) + (n - b) = k$. Using $a + b = k$ provides $k = n$. Thus $a + b = n$, which means that they meet only on the diagonal DD'.

The probability that spider meets fly at $(a, n - a)$, it moves a steps east (hence $n - a$ steps north), so

$$P(X=a) = \binom{n}{a}\left(\frac{1}{2}\right)^{n-a}\left(\frac{1}{2}\right)^{a} = \binom{n}{a}\left(\frac{1}{2}\right)^{n}.$$

The probability that fly meets spider at $(a, n - a)$, it moves $(n - a)$ steps west (hence a steps south), so

$$P(Y = n-a) = \binom{n}{n-a}\left(\frac{1}{2}\right)^{a}\left(\frac{1}{2}\right)^{n-a} = \binom{n}{a}\left(\frac{1}{2}\right)^{n}.$$

$\therefore$ $P\{\text{Insects meet at } (a, n-a)\} = P(X = a,\ Y = n - a)$

$$= P(X=a)\,P(Y=n-a) = \left[\binom{n}{a}\left(\frac{1}{2}\right)^{n}\right]^2$$

$$P(\text{Insects meet}) = \sum_{a=0}^{n} \binom{n}{a}^2 \left(\frac{1}{2}\right)^{2n} = \left(\frac{1}{2}\right)^{2m} \sum_{a=0}^{n} \binom{n}{a}^2 = \binom{2n}{n}\left(\frac{1}{2}\right)^{2n},$$

since $$\binom{n+m}{r} = \sum_{j=0}^{r} \binom{n}{j}\binom{m}{r-j} \Rightarrow \binom{2n}{n} = \sum_{j=0}^{n} \binom{n}{j}\binom{n}{j}, \quad (m = n = r).$$

Problems with Solutions Provided at the End of the Text

1*. In a family of four children, find the probability that there will be (i) at least one boy, (ii) at least one boy and at least one girl. Assume that the probability of a male birth is $p = 1/2$.

2*. A family has 5 children. Find the probability that this family has at least one girl, given that they have at least one boy. Assume that either sex-birth is equiprobable and all 5 births were independent.

3*. A radar system has a probability of 0.1 of detecting a certain target during a single scan. Find the probability that the target will be detected (a) at least two times in four consecutive scans (event A), (b) at least once in twenty scans (event B).

4*. A room has three lamp sockets. From a collection of 10 light bulbs of which only 6 are good, a person selects 3 at random and puts them in the sockets. What is the probability that room will have light ?

5*. A and B play a game in which their chances of winning are in the ratio 3 : 2. Find A's chance of winning at least three games out of five games played.

6*. An experiment succeeds as twice often as it fails. Find the chance that in the next six trials, there shall be at least four successes.

7*. What is the probability of obtaining 50 heads in a throw of 100 coins and then repeating this event two times in the next five throws ?

8*. If a fair coin is flipped an even number $2n$ times, show that the probability of obtaining more heads than tails is

$$\frac{1}{2}\left\{1 - \binom{2n}{n}\left(\frac{1}{2}\right)^{2n}\right\}.$$

9*. Show that if a coin is tossed n times, the probability of not more than r heads is

$$\left[\binom{n}{0} + \binom{n}{1} + \dots + \binom{n}{r}\right]\left(\frac{1}{2}\right)^n.$$

10*. If E_r denotes exactly r failures, find the chance of success if $P(E_r)/P(E_{n-r})$ in n trials is independent of n.

11*. An irregular six-faced die is thrown and the expectation that in 10 throws it will given five even numbers is twice the expectation that it will give four even numbers. How many times in 10,000 sets of 10 throws, would you expect to give no even numbers ?

12*. Suppose n balls are distributed at random into N boxes. Find the probability that there are exactly k balls in the first r boxes.

13*. Of two equally strong ping-pong players A and B, is it more probable that (a) A will beat B in 3 games out of 4, or in 5 games out of 8, (b) at least 3 games out of 4 or at least 5 out of 8 ?

14*. A certain airline sells 100 seats on a plane that has 95 seats, so as to compensate 5% defaulters. Find the chance that there will be a seat available for every person who shows up for the flight.

15*. If m things are distributed among a men and b women, show that the chance that the number of things received by men is odd, is

$$\tfrac{1}{2}[(b+a)^m-(b-a)^m]/(b-a)^m.$$

16*. If a coin is tossed n times, where n is very large even number, show that the probability of getting exactly $(\frac{1}{2}n-x)$ heads and $(\frac{1}{2}n+x)$ tails is approx.

$$(2/\pi n)^{1/2}\exp(-2x^2/n).$$

Exercise 12(a)

1. The probabilities that in a factory, a worker is skilled is 0.4. Find the probability that out of 5 workers, (i) none, (ii) one, (iii) at least one, will be skilled.
2. Ten coins are thrown simultaneously. Find the probability of obtaining at least seven heads. What is the probability of getting 6 or more heads ?
3. Eight coins are thrown simultaneously. Show that the chance of obtaining at least six heads is 37/256.
4. A man tosses a fair coin 10 times. Find the probability that he will have
   (a) heads on the first five tosses and tails on the next five tosses.
   (b) heads on tosses 1, 3, 5, 7, 9 and tails on tosses 2, 4, 6, 8, 10.
   (c) 5 heads and 5 tails ; (d) at least 5 heads; (e) not more than 5 heads.
5. If on average, 1 vessel in every 10 is wrecked, find the chance that out of 5 vessels expected to arrive, at least 4 will arrive safely.
6. In the long run, 3 vessels out of every 100 are sunk. If 10 vessels are out, what is the probability of safe arrival of (a) Exactly 6, (b) At least 6, (c) At most 6.
7. In a 20-question, 5 answer multiple-choice examination, what is the probability of getting 6 or more correct by mere guessing ?
8. If an ordinary die is thrown 4 times, what is the probability that
   (a) exactly two 6's occur ? (b) Two or fewer 6's occur ?
9. Six dice are thrown 729 times. How many times do you expect at least 3 dice to show a five or six ? [$f=233$]
10. If the birth sex ratio is 49 girls and 51 boys, find the probability of there being 8 or more girls amongst 10 babies born on the same day in a maternity hospital.
11. Out of 800 families with 4 children each, how many families would be expected to have (a) 2 boys and 2 girls, (b) at least one boy, (c) no girl, (d) at most 2 girls ?
    Assume equal birth-probabilities for boys and girls.
12. Suppose that the probability is 0.55 for the birth of a male child and that two successive births are independent. If a woman is to have five children, find the probability that she will have :

(a) 3 boys and two girls (b) Son, daughter, son, daughter, son in that order

(c) children of both sexes (d) two sons, given that she already has 2 sons.

13. (a) An ordinary die is thrown 4 times. Find the chances of obtaining 4, 3, 2, 1, 0 aces.

(b) In 4 throws with a pair of dice, find the chance of throwing doublets twice at least.

14. Which of the three events A, B, C is more probable : A — at least one six when 6 dice are cast, B — at least two sixes when 12 dice are cast and C — at least three sixes when 18 dice are cast ?

15. A perfect cubical die is thrown a large number of times in sets of 8. The occurrence of 5 or 6 is called a success. In what proportion of the sets you expect 3 successes ? [Ans. [illegible]]

16. A boy can jump 5 ft. 3 in. high, three times out of five attempts. What is the chance that, in a random sample of his jumps, 3/7 of them are successful ?

17. A throws 3 coins and B throws 2 coins. Find the chance that A will throw greater number of heads than B.

18. The probability of hitting a runaway with a bomb dropped from a bomber is 1/4. Determine the probability of making the runaway inserviceable if at least 4 hits are required for the purpose on the assumption that 8 bombs are dropped in all.

19. Eight mice are selected at random from a large number and then divided into 2 groups of 4 each, group A and group B. Each mouse in group A is given a doze a of a certain poison which is expected to kill 1 in 4. Each mouse in B is given a dose b of another poison which is expected to kill 1 in 2. Show that nevertheless, there may be fewer deaths in group B than in A, and the probability of this particular happening is 525/4096.

20. A person will accept a batch of 100 items if a random sample of 5 items from the lot contains not more than one defective item. Determine the probability that the person will accept the batch if it contains 10 defective items. 59%

21. In a precision bombing attack, there is a 5% chance that any one bomb will strike the target. Two direct hits are required to destroy the target completely. How many bombs must be dropped to give a 99% chance or better, of completely destroying the target ? [$n = 11$]

22. Electric motors are packed in lots of 10. Before exporting, an inspector choose 3 motors without replacement and if none of these is defective, it is passed. If actually there are three defectives in a lot, what is probability that the lot is passed ?

Find also the probabilities that the sample contains 1, 2, 3 defectives. (No simplification is necessary.)

23. A group of n cars enters an intersection from the south. Through prior observations, it is known that each car has the probability p of turning east, prob. q of turning west, and prob. r of going straight. ($p + q + r = 1$) Assume that drivers behave independently and let X be the number of cars turning east Y the number turning west. Determine the joint p.d.f. $P_{X,Y}(x, y)$.

24. One of the numbers 1, 2, ..., 6 is to be chosen by casting an unbiased die. Let this random experiment be repeated five independent times. Let the variate X be the number of terminations in the set $\{x \mid x = 1, 2, 3\}$ and let the variate Y be the terminations in the set $\{x : x = 4, 5\}$. Compute $P(X = 2, Y = 1)$.

25. (a) Let an unbiased die be cast at random seven independent times. Compute the conditional prob. that each side appears at least once relative to the hypothesis that the face 'one' appears exactly twice.

(b) A coin is tossed 30 times. Events A and B are that first 10 tosses resulted in heads and heads appeared at least 10 times. For $1 \le n \le 20$, what is the conditional probability that exactly $(10 + n)$ heads appear, given A. Show that

$$P\{(10+n) \text{ tosses yield heads} \mid B\} = \binom{30}{10+n}\left(\frac{1}{2}\right)^n \sum_{k=0}^{20}\binom{30}{10+k}\left(\frac{1}{2}\right)^k$$

26. An owner of five overnight cabins is considering buying television sets to rent to cabin occupants. He expects that about half of his customers would be willing to rent sets, and finally he buys three sets. Assuming 100% occupancy at all times :

(a) What fraction of the evenings will there be more requests than TV sets ?

(b) What is the prob. that a customer who requests a TV set will get one.

(c) If owner's cost per set per day is c, what rental R must be charged in order to break even (no gain or loss) in the long run ?

27. The chance of success in each Bernoulli trial is p. If p_k is the probability that there are even number of successes in k trials prove that $p_k = p_{k-1}(1-2p) = qp_{k-1} + p(1-p_{k-1})$. Deduce :

$$p_k = \tfrac{1}{2}[1+(1-2p)^k] = \tfrac{1}{2}[1+(q-p)^k].$$

28. Two players A and B, want respectively m and n points in order to win a set of games. If they separate without playing out of the game, how should the stakes be divided ?

29. A and B have equal chances of winning a single game, A requires n games and B, $n+1$ games to win a match. Show that the odds in favour of A are $1+P : 1-P$ where $P = 2n\,!/(n\,!\,2^n)^2$.

30. A group of $2n$ boys and $2n$ girls are divided into two equal groups. Find the probability p such that each group will be equally divided into boys and girls. Estimate p using Stirring's formula for n !

31. [Y.P. Sabharwal] In a sequence of symmetric Bernoulli Trials, let B_n denote the number of trials covering the first and the last success. Show that

$$\mathrm{E}(B_n) = (n-2)+(n+2)2^{-n},\ \mathrm{Var}(B_n) = 4 - n(2n+1)\,2^{-n} - (n+2)^2\,4^{-n}.$$

12-20. *Probability Recurrence Formula*

When computing the binomial probabilities numerically, one can often use a process of iteration based on the chain relation

$$f(x+1) = \frac{n-x}{x+1}\frac{p}{1-p}\cdot f(x), \qquad 0<p<1. \qquad \ldots(1)$$

To prove it, we notice that

$$\frac{f(x+1)}{f(x)} = \binom{n}{x+1} q^{n-x-1}\,p^{x+1} \Big/ \binom{n}{x} q^{n-x} p^x = \frac{n-x}{x+1}\frac{p}{q}. \qquad \ldots(2)$$

Eqn. (2) is equivalent to (1). Further $f(0) = (1-p)^n = q^n$, is first evaluated and then (1) is applied to compute the remaining frequencies $f(x)$, one after the other.

12-21. *Worked-out Problems*

***Example* 1.** An urn contains balls numbered 1, 2, 3. First a ball is drawn from the urn and then a fair coin is tossed the number of times as the number shown on the drawn ball. Find the expected number of heads.

Solution. Let B_n = {ball numbered n is drawn}, n = 1, 2, 3; and X = Number of heads turned up. Since $P(B_n)$ = 1/3 as balls are equally probable hence by Total-E Rule.

$$\therefore \qquad \mathrm{E}(X) = \sum_{n=1}^{3} P(B_n)\mathrm{E}(X\,|\,B_n) = \frac{1}{3}\sum_{n=1}^{3}\mathrm{E}(X\,|\,B_n).$$

Now $(X \mid B_n)$ = No. of heads when it is tossed n times ; so $(X \mid B_n) \sim \text{bin}(n, p)$. As such

$$\text{E}(X \mid B_n) = np = \frac{n}{2}; \text{ so } \text{E}(X) = \frac{1}{6}\sum_{n=1}^{3} n = 1.$$

***Example* 2.** An N-sided die is tossed and if n shows on its top face, a coin with prob p of heads is tossed n times resulting in m heads $(0 \le m \le n)$. Record the sample space and the probability for (n, m).

Solution. Here $S = \{(n, m) : n = 1, 2, 3, ..., N ; 0 \le m \le n\}$; because there can be no fewer than zero heads and no more than n heads on n flips.

Let X denote the outcome of the Die and Y denote the outcome of the coin-toss. Then

$$P\{Y = m \mid X = n\} = P\{\text{coin tossed } n \text{ times results in } m \text{ heads}\} = \binom{n}{m} q^{n-m} p^m$$

$$P(X = n) = P\,(\text{face } n \text{ occurs on } N\text{-sided die}) = 1/N.$$

$$\therefore \quad P\{X = n, Y = m\} = P\,(X = n).\, P\{Y = m \mid X = n\} = \frac{1}{N}\binom{n}{m} q^{n-m} p^m.$$

This is the probab. $f(n, m)$.

Problems with Solutions Provided at the End of the Text

1*. Five balls were drawn, one at a time, with replacement, from a bag containing an equal number of black and white balls. The number of black balls X was then tabulated for 819 sets of consecutive drawings to give the following observed frequency (f) distribution

X :	0	1	2	3	4	5
f :	30	125	277	224	136	27

Fit the binomial distribution.

2*. If on the average, rainfalls on ten days in every thrity days, find the prob. that
(i) rain will fall on at least three days of a given week.
(ii) first four days of a given week will be wet and the remaining dry (fine).

3*. If on an average 9 ships out of 10 arrive safely to a port, obtain mean and S.D. of the number of ships returning safely out of a total of 150 ships.

12-30. *Some Recurrence Relations for bin (n, p)*

Let $X \sim \text{bin}(n, p)$, $Y = (n - X) \sim \text{bin}(n, q)$ and denote as under :

$$B(k\,; n, p) = P\{X \le k\} = \sum_{x=0}^{k} b(x\,; n, p) = \sum_{x=0}^{k} \binom{n}{x} p^x q^{n-x}.$$

1. $b(k\,; n, p) = b(n-k\,; n, q)$ [Valueable for bin. Tables]
2. $b(k\,; n, p) = B(k, n, p) - B(k-1, n, p)$.
3. $\sum_{k=r}^{n} b(k; n, p) = 1 - \sum_{k=m}^{n} b(k; n, q),\ m = n - r + 1.$
4. $b(k; n+1, p) = p\,.\,b(k-1, n, p) + q\,.\,b(k\,; n, p)$.
5. $B(k, n+1, p) = p\,.\,B(k-1, n, p) + q.\,B(k\,; n, p)$.
6. $b(k, n, p) = B(n-k, n, q) - B(n-k-1\,; n, q)$.

Proof. **1.** $b(k\,;n,p) = \binom{n}{k} p^k q^{n-k} = \binom{n}{n-k} q^{n-k} p^k = b(n-k\,;n,q).$

2. $P(X \le k) - P(X \le k-1) = P(X = k)$ yields the result.

3. $P(X \ge r) = 1 - P(X \le r-1) = 1 - P(n - X \ge n - r + 1) = 1 - P(Y \ge m).$

This amounts to stated result.

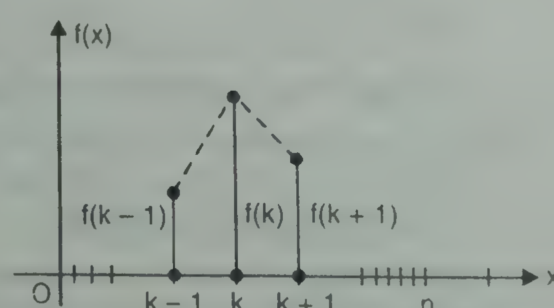
4.
$$\text{R.H.S.} = p\binom{n}{k-1} p^{k-1} q^{n+1-k} + q\binom{n}{k} p^k q^{n-k}$$
$$= \left[\binom{n}{k-1} + \binom{n}{k}\right] p^k q^{n+1-k} = \binom{n+1}{k} q^{n+1-k} p^k = b(k, n+1, p).$$

5. From Iteration formula in (4) we get
$$b(x\,;n+1,p) = pb(x-1,n,p) + q\,.\,b(x\,;n,p)$$
Summing up from $x = 0$ to $x = k$, this provides
$$B(k\,;n+1,p) = pB(k-1,n,p) + q\,B(k;n,p).$$

6. $\text{R.H.S.} = \sum_{x=0}^{n-k} \binom{n}{x} q^x p^{n-x} - \sum_{x=0}^{n-k} \binom{n}{x} q^x p^{n-x} = \binom{n}{n-k} q^{n-k} p^k = \binom{n}{k} p^k q^{n-k} = \text{L.H.S.}$

12-31. *Mode of the Binomial Distribution*

We have to find the value of the random variable $X \sim \text{bin}(n,p)$ for which the probability is maximum. Thus, if k is the modal value, its definition provides : $f(k) \ge \max\ \{f(k+1), f(k-1)\}$.



The probability recurrence formula [§ 12-20] gives

$$\frac{f(k)}{f(k+1)} = \frac{(k+1)\,q}{(n-k)\,p};\quad \frac{f(k)}{f(k-1)} = \frac{(n-k+1)\,p}{kq}.$$

$\therefore\ f(k) \ge f(k+1) \Rightarrow k > np - q = (n+1)\,p - 1;\quad f(k) \ge f(k-1) \Rightarrow k \le (n+1)\,p.$

Combining $\boxed{(n+1)\,p - 1 \le k \le (n+1)\,p.}$ ...(1)

Now two cases arise :

***Case* 1.** Let $(n+1)p$ be an integer. Since the integer k lies between two consecutive integers $(n+1)p$ and $(n+1)p-1$, equation (1) is possible iff $k = (n+1)p$ or $k = (n+1)p - 1$. Thus there are two modes : $(n+1)p$ and $(n+1)p - 1$. And the distribution is then said to be bimodal.

***Case* 2.** Let $(n+1)\,p$ be a fraction (rational number). Then integer k lies between two fractions differing by 1. Hence k must be integral part of $(n+1)p$.

Remark. The conclusions from (1) can be most easily grapsed by assuming some concrete values for $(n+1)p$.

Example : Show that if np be a whole number, the mean of the binomial coincides with the greatest term.

Solution. Since np is a whole number and $0 < p < 1$, we observe that $(n + 1)p$ is a fraction. Also mode of the binomial distribution is the integral part of $(n + 1)p$, hence

$$np = [(n+1)\,p] \Rightarrow \text{mean} = \text{mode}.$$

Since the greatest term of the binomial $(q + p)^n$ is the modal value, hence the mean coincides with the greatest term (mode) in case np is a whole number.

Problems with Solutions Provided at the End of the Text

1*. Compute the mode of a binomial distribution with (i) $n = 7, p = 1/2$, (ii) $n = 50, p = 1/16$.

2*. A pays Rs 1 for each participation in the following game. Three dice are tossed ; if one six appears he gets Rs 1; if two sixes appear he gets Rs 2 and if three sixes appear he gets Rs 8; otherwise he gets nothing. Is the game fair ? If not, how much should the player get when three sixes appear to make the game fair ?

3*. Let $X \sim \text{bin}\,(n, p)$. Use Chebyshev's Inequality to find the bounds when

(a) $n = 40, p = 1/4$; $P\{\,|X - 10| < 8\}$, $P\{\,|X - 10| > 10\}$,

(b) $n = 40, p = 1/2$; $P\{\,|X - 20| \leq 5\}$,

(c) $n = 40, p = 1/2$; $P\{\,|X - 20| > 10\,k\} \leq 0.25$, find k.

Exercise 12(b)

1. Find the most probable number of heads in 99 tossings and 100 tossings of a biased coin given that the probability of a head in a single tossing is 3/5.
2. For a certain basket ball player, the probability of throwing the ball into the basket in one throw is 0.4. He makes 10 throws. Find the most probable number of successful throws and the corresponding probability.
3. Find the most probable number of negative and positive errors and the corresponding probabilities in four measurements if, in each of them, the probability of a positive error equals 2/3 and of a negative error equals 1/3.
4. Let X be bin (n, p). If $0 < p < (n + 1)^{-1}$, then the biggest probability is at the possible value 0 and the successive probabilities are strictly decreasing. If $(n + 1)^{-1} < p < 1/2$, then the distribution is symmetric, skewed to the left. It is strictly increasing until it achieves a maximum.
5. The following data due to Weldon shows the results of throwing 12 dice 4096 times, a throw of 4, 5 or 6 being called a success,

x :	0	1	2	3	4	5	6	7	8	9	10	11	12	Total
f :	–	7	60	198	430	731	948	847	536	257	71	11	–	4096

Fit the binomial distribution and calculate the expected frequencies. Campare the actual mean and S.D. with those of the expected ones for the distribution.

6. A set of six similar coins is tossed 640 times with the following results :

Number of heads :	0	1	2	3	4	5	6
Frequency :	7	64	140	210	132	75	12

Calculate the binomial frequencies on the assumption that the coins are symmetrical.

7. Seven coins are tossed and the number of heads noted. The experiment is repeated 128 times with the following results :

Number of heads	: 0	1	2	3	4	5	6	7
Frequency	: 7	6	19	35	30	23	7	1

Fit the binomial distribution assuming :

(a) the coin is unbiased (b) the nature of the coin is unknown

(c) the p for 4 coins is 0.50 and for other 3 is 0.45.

8. The distribution of headless matches per box of 50 in a total of 100 boxes is as under :

No. of headless matches per box :	0	1	2	3	4	5	6	7
No. of boxes :	12	27	29	19	8	4	1	0

Assuming the distribution of headless matches per box over 100 boxes is binomial, fit a binomial curve to the above data and test the goodness of the fit. Compare the variance of the actual and observed distribution. [**Hint.** $\bar{x} = 2$, $\Sigma f(x-\bar{x})2/N = 1.78$, $npq = 1.92$]

9. In 103 litters of 4 mice the number of litters which contained 0, 1, 2, 3, 4 females were noted as under :

No. of female mice :	0	1	2	3	4	Total
No. of litters :	8	32	34	24	5	103

(a) If the chance of obtaining a female in a single trial is supposed constant, estimate this unknown constant probability.

(b) If the size of the litter (4) was not given, how could it be estimated from the data ?

(c) How could the assumption that the chance of obtaining a female in a single trial is constant, be tested ?

10. The screws produced by a certain machine were checked by examining a sample of 12. The following table shows the distribution of 50 samples according to the number of defective items they contained.

x :	0	1	2	3	4	5 or more	Total
f :	18	19	9	3	1	0	50

[X = No. of defective in a sample of 12, f = No. of samples]. Calculate the mean number of defectives per sample and assuming that the binomial law holds, estimate the proportion of defective items in the whole lot produced by the machine.

11. Define Bernoulli trials. Find the bounds for the most probable number of successes in a sequence of Bernoulli trials.

One worker can manufacture 120 articles during a shift, another worker 140 articles, the chances of the articles being of a high quality are 0.94 and 0.80 respectively. Determine the most probable number of high quality articles manufactured by each worker.

12-40. *Simple Factorial Moments of bin (n, p)*

$$\mu'_{(r)} = E(X^{(r)}) = \sum_{x=0}^{n} \binom{n}{x} q^{n-x} p^x x^{(r)} \qquad \left[\binom{n}{x} = \frac{n^{(r)}}{x^{(r)}}\binom{n-r}{x-r}\right] \qquad \ldots(1)$$

$$= \sum_{x=r}^{n} n^{(r)} \binom{n-r}{x-r} q^{n-x} p^x = n^{(r)} p^r \sum_{x=r}^{n} \binom{n-r}{x-r} q^{n-x} p^{x-r} = n^{(r)} p^r \,.\, (q+p)^{n-r} = n^{(r)} p^r.$$

Thus $$\boxed{\mu'_{(r)} = n^{(r)} p^r.}$$

Note. Since $\binom{X}{r} = \dfrac{X^{(r)}}{r!}$, so, $E\binom{X}{r} = \dfrac{E[X^{(r)}]}{r!} = \binom{n}{r}p^r$.

P.G.F. Method. $G(t) = (q+pt)^n = \sum\limits_{x=0}^{n}\binom{n}{x}q^{n-x}p^x t^x$ [Binomial expansion]

We differentiate this identity r times w.r.t. 't' to get

$$n^{(r)}p^r(q+pt)^{n-r} = \sum_{x=0}^{n}\binom{n}{x}q^{n-x}p^x x^{(r)} t^{x-r}$$

$$\therefore \quad n^{(r)}p^r = \sum_{x=0}^{n}\binom{n}{x}q^{n-x}p^x x^{(r)} = E(X^{(r)}) \qquad [\text{Put } t = 1,\ p+q=1]$$

Simple Moments through Simple Factorial Moments

$$\underline{\mu_2'} = E(X^2) = E(X^{(2)} + X) = \mu_{(2)}' + \mu_{(1)}' = \underline{n^{(2)}p^2 + np.}$$

$$\underline{\mu_3'} = E(X^3) = E(X^{(3)} + 3X^{(2)} + X) = \mu_{(3)}' + \mu_{(2)}' + \mu_{(1)}' = \underline{n^{(3)}p^3 + 3n^{(2)}p^2 + np.}$$

$$\underline{\mu_4'} = E(X^4) = E(X^{(4)} + 6X^{(3)} + 7X^{(2)} + X)$$

$$= \mu_{(4)}' + 6\mu_{(3)}' + 7\mu_{(2)}' + \mu_{(1)}' = \underline{n^{(4)}p^4 + 6n^{(3)}p^3 + 7n^{(2)}p^2 + np.}$$

12-41. *Central Moments through Simple Moments*

$$\underline{\mu_2} = \mu_2' - \mu_1'^2 = [n(n-1)p^2 + np] - (np)^2 = np - np^2 = \underline{npq.}$$

$$\underline{\mu_3} = \mu_3' - 3\mu_2'\mu_1' + 2(\mu_1')^3$$

$$= [n(n-1)(n-2)p^3 + 3n(n-1)p^2 + np] - 3[n(n-1)p^2 + np]np + 2(np)^3$$

$$= (p^3 - 3p^3 + 2p^3)n^3 + (-3p^3 + 3p^2 + 3p^3 - 3p^2)n^2 + (2p^3 - 3p^2 + p)n$$

$$= np(1 - 3p + 2p^3) = np(1-p)(1-2p) = \underline{npq(q-p).}$$

$$\underline{\mu_4} = \mu_4' - 4\mu_3'\mu_1' + 6\mu_2'\mu_1'^2 - \mu_1'^4$$

$$= [n^{(4)}p^4 + 6n^{(3)}p^3 + 7n^{(2)}p^2 + np] - 4np[n^{(3)}p^{(3)} + 3n^{(2)}p^2 + np] + 6(np)^2[n^{(2)}p^2 + np] - 3(np)^4.$$

We expand $n^{(4)}$, $n^{(3)}$, etc. and collect the coefficients, thus

Coeff. of $n^4 = p^4 - 4p^4 + 6p^4 - 3p^4 = 0.$

Coeff. of $n^3 = -6p^4 + 6p^3 + 12p^4 - 12p^3 + 6p^3q = 0.$

Coeff. of $n^2 = 11p^4 - 18p^3 + 7p^2 - 4p^2 + 12p^3 - 8p^4 = 3p^2(1 - 2p + p^2) = 3p^2q^2.$

Coeff. of $n = -6p^4 + 12p^3 - 7p^2 + p = -6p^2(p^2 - 2p + 1) - p^2 + p = -6p^2q^2 + pq = pq(1 - 6pq)$

Hence $\mu_4 = 3n^2p^2q^2 + npq(1 - pq) = \underline{npq(1 + 3npq - 6pq).}$

12-42. *Central Factorial Moments*

$\underline{\mu_{(1)}} = \mu_1' = \mu = np;\quad \mu_{(2)} = \mu_2 = \underline{npq}$

$\mu_{(3)} = \mu_3 - 3\mu_2 = npq(q-p) - 3npq = -npq[3(p+q) - (q-p)] = -2npq(2p+q) = -2pq(1+p).$

$$\mu_{(4)} = \mu_4 - 6\mu_3 + 11\mu_2 = [3n^2p^2q^2 + npq(1-6pq)] - 6[npq(q-p) + 11npq]$$
$$= npq[3npq + 1 - 6pq - 6q + 6p + 11] = 3npq[npq + 4 - 2p(1-p) - 2(1-p) + 2p]$$
$$= \underline{3npq[npq + 2(1+p+p^2)]}.$$

12-43. Worked-out Problems

***Example* 1.** Let $X \sim$ bin (n, p) and $a > 0$. Prove that

$$P\left\{\left|\frac{X}{n} - p\right| > a\right\} \le \frac{\sqrt{pq}}{na^2} \min\left\{a\sqrt{n}, \sqrt{pq}\right\}.$$

Solution. Let $Z = X - np$, then $\mathsf{E}(Z) = 0$, Var $(Z) = npq$ and

$$\mathsf{E}\,|Z| \le \sqrt{\mathsf{E}(|Z|^2)} = \sqrt{\mathsf{E}(X-np)^2} = \sqrt{\text{Var}(X)} = \sqrt{npq}$$

where the first inequality is Cauchy-Schwarz. Now

$$p_1 = P\left\{\left|\frac{X}{n} - p\right| > a\right\} = P\{|Z| > na\} \le \mathsf{E}(|Z|)/na \le \sqrt{npq}/na \qquad \text{[Markov Inequality]}$$

$$p_1 = P\{|Z| > na\} \le \text{Var}(X)/n^2a^2 = npq/n^2a^2, \qquad \text{[by Chebyshev's Inequality]}$$

Thus, $\quad p_1 \le \min\{\sqrt{pq}/a\sqrt{n},\ pq/na^2\}, \qquad$ [Composite result]

$$= (\sqrt{pq}/na^2)\min\{a\sqrt{n}, \sqrt{pq}\}.$$

***Example* 2.** If X is bin (n, p), show that

$$\mathsf{E}\left(\frac{X}{n} - p\right)^2 = \frac{pq}{n}, \qquad \text{Cov}\left(\frac{X}{n}, \frac{n-X}{n}\right) = -\frac{pq}{n}\ \text{Cov}(X, n-X) = -npq.$$

Solution. $\mathsf{E}\left(\dfrac{X}{n} - p\right)^2 = \dfrac{1}{n^2}\mathsf{E}(X-np)^2 = \dfrac{\text{Var}(X)}{n^2} = \dfrac{npq}{n^2} = \dfrac{pq}{n}.$

Recall : Cov $(aX + b, cY + d) = ac$ Cov (X, Y).

$$\therefore \quad \text{Cov}\left(\frac{1}{n}X, \frac{1}{n}X + 1\right) = -\frac{1}{n^2}\text{Cov}(X, Y) = -\frac{1}{n^2}\text{Var}(X) = -\frac{npq}{n^2} = -\frac{pq}{n}$$

$$\text{Cov}(X, n-X) = -\text{Cov}(X, X) = -\text{Var}(X) = -npq.$$

***Example* 3.** Show that the necessary and sufficient conditions for two given numbers, a, b to be mean and variance of some binomial distribution are that $a > b > 0$ and $a^2/(a-b)$ is an integer. Show further that when these conditions are satisfied, the binomial distribution is uniquely determined.

Solution. *Necessity :* Let $a = np$, $b = npq$, then $a > b > 0$ and

$$a^2/(a-b) = n^2p^2/np(1-q) = n \text{ (positive integer)}.$$

Sufficiency : Let $a > b > 0$ and $a^2/(a-b) = n$ (positive integer)

Put $b = af$, where $0 < f < 1$. Now

$$a^2/(a-b) = a^2/a(1-f) = a/(1-f) = n \qquad \text{[by hypothesis]}$$

Thus $a = n(1-f)$ and $b = nf(1-f)$. So, if we choose $f = q$, then $a = np$ and $b = npq$ are the mean and variance of bin (n, p).

Since bin (n, p) is uniquely determined by n and p which in turn uniquely determine a and b, the assertion is proved.

***Example* 4.** It is known that on the average 2/3 of the seeds of a certain variety programme germinate. Use Chebyshev's inequality to obtain an upper bound on the probability that the number germinating will differ from the expected number by more than 10 if 100 seeds are planted.

Solution. Here, $n = 100$, $p = 2/3$, $q = 1/3$, so that $\mu = np = 200/3$, $\sigma^2 = npq = 200/9$. Now Chebyshev's inequality provides :

$$P\{X-\mu \mid \geq c\} \leq \sigma^2/c^2 \;\Rightarrow\; P\{|X-\mu| \geq 10\} \leq 200/900 = 2/9.$$

Thus, the max prob. = 2/9 so that the upper bound to the probability is 2/9.

***Example* 5.** Let $X \sim$ bin (n, p). Show that

$$P\{X - np > n\varepsilon\} \leq \mathrm{E}\{e^{\lambda(X-np-n\varepsilon)}\}, \quad \lambda > 0,\ \varepsilon > 0 \qquad \text{...(1)}$$

Deduce : $P\{|X - np| \leq n\varepsilon\} \to 1$, as $n \to \infty$ ...(2)

Assume :
$$0 < 1 < e^x \leq x + e^{x^2}. \qquad \text{...(3)}$$

Solution. Write $P(X = k) = p_k$. For $k > np + n\varepsilon$, using $p + \varepsilon = \theta$, we have

$$\sum_{k>n\theta} p_k < \sum_{k>n\theta} e^{\lambda(k-n\theta)}\, p_k < \sum_{k=0}^{n} e^{\lambda(k-n\theta)}\, p_k \qquad [1 < e^y]$$

$$= e^{-\lambda n\theta}\left\{\sum_{k=0}^{n} p_k\, e^{\lambda k}\right\}$$

$$= e^{-\lambda n\theta}.(q + pe^{\lambda})^n. \qquad \text{[by m.g.f. of bin } (n, p)] \qquad \text{...(4)}$$

$$= e^{-\lambda n\varepsilon}.\{e^{-\lambda p}(q + pe^{\lambda})\}^n = e^{-\lambda n\varepsilon}\{pe^{\lambda q} + qe^{-\lambda p}\}^n \qquad \text{...(i)}$$

We now utilize extreme right of (3)

$$pe^{\lambda q} + qe^{-\lambda p} \leq p(\lambda q + e^{\lambda^2 q^2}) + q(-\lambda p + e^{\lambda^2 p^2}) = pe^{\lambda^2 q^2} + qe^{\lambda^2 p^2}$$

$$\leq (p+q)\,e^{\lambda^2}, \quad [\lambda^2 p^2 < \lambda^2, \lambda^2 q^2 < \lambda^2] \qquad \text{...(ii)}$$

Substituting from (ii) into (i) and noting that L.H.S. of (4) is $P\{X > n\theta\}$, we get

$$P\{X > n\theta\} \leq e^{-\lambda n\varepsilon + n\lambda^2} = e^{-n\varepsilon^2/4}. \qquad [\text{Choose } \lambda = \varepsilon/2] \qquad \text{...(iii)}$$

i.e.
$$P\{(X - np) > n\varepsilon\} \leq e^{-n\varepsilon^2/4}$$

Similarly : $P\{X - np < -n\varepsilon\} \leq e^{-n\varepsilon^2/4}$

$\therefore$
$$P\{|X - np| > n\varepsilon\} \leq 2\,e^{-n\varepsilon^2/4} \to 0, \text{ as } n \to \infty. \qquad \text{...(iv)}$$

i.e.
$$P\{|X - np| \leq n\varepsilon\} \to 1, \text{ as } n \to \infty \qquad \text{[by Neg-Rule]}$$

Note that the R.H.S. of (1) and (4) are identical.

Problems with Solutions Provided at the End of the Text

1*. Bring out the fallacy, if any, in the following statement : "The mean of a binomial distribution is 5 and its S.D. is 3".

2*. Show that the largest possible value of the variance of a binomial variate is $n/4$.

3*. For a binomial experiment, show that if μ is near zero, $P(X \geq 1) = \mu$.

4*. For given values of x and n, consider $b\ (x\ ;\ n,\ p)$ as a function of p and show that it is maximized where $p = (x/n)$.

5*. If X is bin $(4, p)$, find $\mathsf{E}\ [\sin\ (\pi X/2)]$.

6*. For a binomial distribution (B.D.), the mean is 6 and S.D. is $\sqrt{2}$. Write out all the terms of B.D.

7*. If X is binomially distributed with $\mu = 4$ and $\mu_3 = 1.92$, find the other constants of the distribution.

12-50. *Mean (absolute) Deviation*

For bin (n, p), mean $= np$, so $M = \mathsf{E}\ (\ |\ X - np\ |)$.

$$\therefore \qquad M = \sum_{x=0}^{n} f(x)\,|\,x - np\,| = \sum_{x<np} f(x)\,(np - x) + \sum_{x>np} f(x)\,(x - np).$$

$$0 \equiv \mathsf{E}\,(X - np) = \sum_{x=0}^{n} f(x)\,(x - np) = \sum_{x<np} f(x)\,(x - np) + \sum_{x>np} f(x)\,(x - np)$$

We add these equations; write $\lambda = [np + 1]$ for the *greatest integer* function, and obtain

$$M = 2\sum_{x=\lambda}^{n} f(x)\,(x - np), \qquad \text{(use } x = xp + xq\text{)}$$

$$\frac{1}{2}M = \sum_{\lambda}^{n} \frac{n!\,p^x q^{n-x}}{x!\,(n-x)!}\,[qx - (n-x)\,p] = \sum_{\lambda}^{n}\left[\frac{n!\,p^x\,q^{n-x+1}}{(n-x)!\,(x-1)!} - \frac{n!\,p^{x+1} q^{n-x}}{(n-x-1)!\,x!}\right]$$

$$= \sum_{\lambda}^{n}(A_x - A_{x+1}), \text{ where } A_x = (npq)\binom{n-1}{x-1} q^{n-x} p^{x-1}.$$

This finite series telescopes to sum A_λ $[\because A_{n+1} = 0]$.

$$\therefore \qquad \underline{M = (2npq)\binom{n-1}{\lambda-1} q^{n-\lambda} p^{\lambda-1},\ \lambda = [np + 1].}$$

12-60. *M.G.F. and Ch. Function*

$$\underline{M(t : x) = (q + pe^t)^n\ ;\ \ \phi(t : X) = (q + pe^{it})^n} \qquad \text{[See § 8-16]}$$

$$\therefore \qquad \mu = [M'(t)]_{t=0} = np[(q + pe^t)^{n-1} e^t]_{t=0} = np\,(q + p)^{n-1} = np.$$

12-61. *Central Moments through Central M.G.F.*

$$M(t : X - np) = e^{-npt} M(t : X) = e^{-npt}(q + pe^t)^n = (qe^{-pt} + pe^{qt})^n$$

Using exponential series $e^{\theta} = \Sigma(\theta^k / k!)$, this reduces to*

$$M(t : X - np) = \left\{1 + pq\frac{t^2}{2!} + pq\,(q-p)\frac{t^3}{3!} + pq\,(1-3pq)\frac{t^4}{4!} + \ldots\right\}^n \qquad \ldots\text{(i)}$$

We connect it with c.g.f. $K(t)$. Thus

$$K(t) = \ell n\, M(t : X) = \ell n \{M(t : X - np) . e^{npt}\} = npt + \ell n\, M(t : X - np)$$

Using (i), this becomes

$$K(t) = npt + n\,\ell n\left\{1 + pq\frac{t^2}{2!} + pq\,(q-p)\frac{t^3}{3!} + pq\,(1-3pq)\frac{t^4}{4!} + \ldots\right\}$$

We now use logarithmic expansion $\ell n(1+T) = T - (T^2/2) + (T^3/3) -$

$$\sum_1^{\infty} k_r \frac{t^r}{r!} = npt + n\left\{pq\frac{t^2}{2!} + pq\,(q-p)\frac{t^3}{3!} + pq\,(1-3pq)\frac{t^4}{4!} + \ldots\right\} - \frac{n}{2}\left\{pq\frac{t^2}{2!} + \ldots\right\}^2 + \ldots$$

Equating various coeffts. of $(t)^r$, r = 1, 2, 3, 4, ...

$$k_1 = np,\ k_2 = npq,\ k_3 = npq(q-p),\ k_4 = npq\,(1-6pq), \ldots$$

Recall : $\mu = k_1$, $\mu_2 = k_2$, $\mu_3 = k_3$, $\mu_4 = k_4 + 3k_2^2 = kpq\ (1 - 6pq + 3npq)$, ...

Standardized binomial variate : $X^* = (X - np)/\sqrt{npq}$.

A standardized binomial variate itself is never a binomial variate.

12-62. *Romanvosky Moments Recurrence Formula*

$$\mu_{r+1} = pq[nr\mu_{r-1} + (d\mu_r / dp)].$$

Proof. By definition, $\mu_r = \mathrm{E}(X - np)^r$

$$\therefore\ \mu_r = \sum_{x=0}^{n} \binom{n}{x} q^{n-x} p^x (x - np)^r, \qquad \left[f(x) = \binom{n}{x} q^{n-x} p^x\right]$$

$$\frac{d\mu_r}{dp} = \sum_{x=0}^{n} \binom{n}{x} \frac{d}{dp}[(p^x q^{n-x}\,(x-np)^r]$$

$$= \sum_{x=0}^{n} \binom{n}{x} \{[xp^{x-1}q^{n-x} - (n-x)\,p^x q^{n-x-1}](x-np)^r - nrp^x q^{n-x}(x-np)^{r-1}\}$$

$$*\ qe^{-pt} + pe^{qt} = q\left[1 - pt + \frac{p^2t^2}{2!} - \frac{p^3t^3}{3!} + \ldots\right] + p\left[1 + qt + \frac{q^2t^2}{2!} + \frac{q^3t^3}{3!} + \ldots\right]$$

$$\text{Coeff.}\frac{t^4}{4!} = qp(p^3 + q^3) = qp\{(p+q)^3 - 3pq(p+q) = pq\,(1-3pq).$$

$$= \sum_{x=0}^{n} \binom{n}{x} p^x q^{n-x} \left\{ \left[\frac{x}{p} - \frac{n-x}{q} \right] (x-np)^r - nr(x-np)^{r-1} \right\}$$

$$= \sum_{x=0}^{n} \frac{f(x)(x-np)^{r+1}}{pq} - nr \sum_{x=0}^{n} f(x)(x-np)^{r-1},$$

$$= (pq)^{-1} \mathsf{E}(X-np)^{r+1} - nr\mathsf{E}(X-np)^{r-1} = (pq)^{-1} \mu_{r+1} - nr\,\mu_{r-1}.$$

Thus, by transfer, $\mu_{r+1} = pq[nr\mu_{r-1} + (d\mu_r/dp)]$.

Note. Put n = 1, 2, 3, in succession, use $\mu_0 \equiv 1$, $\mu_1 \equiv 0$ to obtain

$$\mu_2 = pq\,[n\mu_0 + D\mu_1] = npq. \quad [D = d/dp]$$

$$\mu_3 = pq\,[2n\mu_1 + D\mu_2[= pq\,D(npq) = npq\,(q-p).$$

$$\mu_4 = pq\,[3n\mu_2 + D\mu_3] = 3n^2p^2q^2 + npq\,(1-6pq).$$

***Example* 1.** Suppose that $M(t : X) = (q + pe^t)^n$, x = 0, 1, 2, ..., n. Find the p.m.f. of X.

Solution. $M(t : X) = \sum_{x=0}^{n} \binom{n}{x} q^{n-x} (pe^t)^x = \sum_{x=0}^{n} \binom{n}{x} q^{n-x} p^x e^{xt}$. ...(1)

However, the definition of m.g.f. provides $M(t : X) = \sum_{x=0}^{n} e^{xt} f(x)$.. [Dirichlet's Expansion] ...(2)

Equating the general terms of (1) and (2) we obtain

$$f(x) = {}^nC_x q^{n-x} p^x, \quad x = 0, 1, 2, ..., n, \text{ i.e. } X \sim \text{bin}\,(n, p).$$

***Example* 2.** Using central m.g.f. of bin (n, p), prove that

$$\partial M/\partial t = pq\,\{nt\,M + (\partial M/\partial p)], \qquad [M = M(p, t)] \quad ...(1)$$

Deduce : $\mu_{r+1} = pq\,\{nr\,\mu_{r-1} + (d\mu_r/dp)\}$.

Solution. $M = \mathsf{E}[e^{t(X-np)}] = e^{-npt}(q + pe^t)^n = (pe^{qt} + qe^{-pt})^n = (T)^n$, $\quad [T = pe^{qt} + qe^{-pt}]$

$$\partial M/\partial t = nT^{n-1}.\,pq[e^{qt} - e^{-pt}]$$

$$\partial M/\partial p = nT^{n-1}[e^{qt} - e^{-pt} - t(pe^{qt} + qe^{-pt})] = nT^{n-1}(e^{qt} - e^{-pt}) - ntT^n$$

$$= (pq)^{-1}\,\partial M/\partial t - nt\,T^n$$

This yields : $\dfrac{\partial M}{\partial t} = pq\left\{\dfrac{\partial M}{\partial p} + ntM\right\}$. ...(1)

(ii) We differentiate (1) r times w.r.t. parameter t and use Leibnitz theorem of repeated differentiation to obtain

$$\frac{\partial^{r+1} M}{\partial t^{r+1}} = pq\left\{\frac{\partial}{\partial p}\frac{\partial^r M}{\partial t^r} + n\left[t\frac{\partial^r M}{\partial t^r} + \binom{r}{1}1.\frac{\partial^{r-1} M}{\partial t^{r-1}}\right]\right\}.$$

Put t = 0, use $\partial^k M(0)/\partial t^k = \mu_k$ to set the last result as

$$\mu_{r+1} = pq\,\{(d\mu_r/dp) + nr\mu_{r-1}\}. \quad ...(1)$$

12-63. Cumulant Recurrence Formula

$$\boxed{k_{r+1} = pq(\partial k_r/\partial p).} \qquad \ldots(1)$$

Def. $\quad K(t) = \ln_e M(t) = n\ln(q + pe^t).$

$$\therefore \quad \frac{\partial K}{\partial t} = \frac{npe^t}{q+pe^t}, \quad \frac{\partial K}{\partial p} = \frac{n(e^t-1)}{q+pe^t} \quad \Rightarrow \quad np + pq\frac{\partial K}{\partial p} = \frac{\partial K}{\partial t}.$$

Substituting Taylor's series for function $K(t)$, viz. $\Sigma\kappa_r(t^r/r!)$, this gives

$$np + pq\left\{\sum_1^\infty \frac{\partial \kappa_r}{\partial p}\frac{t^r}{r!}\right\} = \sum_1^\infty \kappa_r \frac{t^{r-1}}{(r-1)!} \qquad \ldots(2)$$

Equating constant term on both sides of (2) gives $\kappa_1 = np$. Next we equate Coeff. of $t^r/(r!)$ on both sides of (1) to obtain

$$\kappa_{r+1} = pq\frac{\partial \kappa_r}{\partial p}.$$

$\kappa_2 = pq\,D(np) = npq$, $\kappa_3 = pqD\,(npq) = npq\,(q-p)$; $\kappa_4 = pq\,D\,\{npq\,(q-p)] = npq\,(1-6pq)$.

$$\beta_1 = \frac{\kappa_3^2}{\kappa_2^3} = \frac{(q-p)^2}{npq}, \qquad \beta_2 = \frac{\kappa_4}{\kappa_2^2} + 3 = \frac{1-6pq}{npq} + 3.$$

$$\gamma_1 = \sqrt{\beta_1} = (q-p)/\sqrt{npq}, \qquad \gamma_2 = \beta_2 - 3 = (1-6pq)/npq.$$

It may be noted that the point (β_1, β_2) lies on the straight line $\beta_1 - \beta_2 = 3 - (2/n)$.

***Example* 1.** If $K(t)$ is the cumulant function of bin (n, p), show that

$$\frac{d}{dt}K(t) = n[1 + e^{-(z+t)}]^{-1}, \text{ where } z = \ln_e(p/q).$$

Further expanding the R.H.S. in powers of t by Taylor's Theorem, show that

$$\kappa_r = nd^{r-1}p/dz^{r-1} \equiv n\,D^{r-1}(p).$$

Hence or otherwise obtain the recurrence relations

$$\kappa_{r+1} = pq\,(d\kappa_r/dp) = (dK_r/dz), \ (r > 1).$$

Solution. By definition. $K(t) = \ln M(t) = \ln(q + pe^t)^n = n\ln q + n\ln\,[1 + e^t(p/q)]$

$$= n\ln q + n\ln[1 + e^{t+z}], \qquad (p/q = e^z, \text{ by hypothesis})$$

$$\text{Thus} \quad \frac{1}{n}\frac{dK}{dt} = \frac{e^{t+z}}{1+e^{t+z}} = \{1 + e^{-(t+z)}\}^{-1} \qquad \ldots(1)$$

Further Parts. Let $\phi(z+t) = [1 + e^{-(t+z)}]^{-1}$, so that $\phi(z) = (1+e^{-z})^{-1} = [1 + (q/p)]^{-1} = p$.

Now Taylor's series for $\phi(z+t)$ in powers of t is, by (1)

$$\frac{1}{n}\frac{dK}{dt} = \phi(z+t) = \sum_{r=0}^\infty \frac{t^r}{r!}\left[\frac{d^r\phi(z)}{dz^r}\right] = \sum_0^\infty \frac{t^r}{r!}\frac{d^r p}{dz^r}. \qquad \ldots(2)$$

Differentiating $K(t) = \Sigma (t^r/r!) k_r$, we get $(dK/dt) = \Sigma k_r t^{r-1}/(r-1)!$

This reduces (2) to the form

$$\sum_{r=1}^{\infty} k_r \frac{t^{r-1}}{(r-1)!} = n \sum_{r=1}^{\infty} \frac{t^{r-1}}{(r-1)!} \frac{d^{r-1}(p)}{dz^{r-1}}$$

So $\qquad k_1 = np, \qquad k_r = n\dfrac{d^{r-1}p}{dz^{r-1}}.$ ...(3)

Finally from (3b),

$$\frac{dk_r}{dp} = n\frac{d}{dz}\left(\frac{d^{r-1}p}{dz^{r-1}}\right).\frac{dz}{dp} \qquad \left[z = \ln\frac{p}{q},\ \frac{dz}{dp} = \frac{1}{p} + \frac{1}{q} = \frac{1}{pq}\right]$$

$$= \frac{n}{pq}\left(\frac{d^r p}{dz^r}\right) = \frac{n}{pq}\left(\frac{k_{r+1}}{n}\right), \qquad \text{[Change } r \text{ to } (r+1) \text{ in (3)]}$$

Thus, $\qquad k_{r+1} = pq\,(dk_r / dp).$

***Example* 2.** In a sequence of n independent Bernoulli trials, let p_n be the probability of an even number of successes. Prove the recurrence formula :

$$p_n = qp_{n-1} + (1 - p_{n-1})p = p + (q - p)\, p_{n-1}$$

where p is the constant probab. of success in each trial. Obtain also the generating function.

Solution. Let S_n^e = {Even number of successes in n trials}, S_n^o = {Odd number of successes in n trials}. T_1 = {First trial is a success}. Now

$S_n^e = T_1 S_{n-1}^o \uplus \bar{T}_1 S_{n-1}^e$ [Success followed by odd no. or failure followed by even}

By Multistage p-Rule

$$P(S_n^e) = P(T_1)\, P(S_{n-1}^o | T_1) + P(\bar{T}_1)\, P(S_{n-1}^e | \bar{T}_1)$$

$$\therefore \qquad p_n = p(1 - p_{n-1}) + q(p_{n-1}) = p + (q - p)\, p_{n-1}.$$

To solve it, we put this in the particular form

$$p_n - \frac{1}{2} = (q - p)\left(p_{n-1} - \frac{1}{2}\right), \quad n \geq 1, \ p_s = 1$$

So $P_n = \lambda P_{n-1}$, $\qquad [\lambda = q - p, P_n = p_n - \frac{1}{2}]$

Obvisouly, $\qquad P_n = \lambda P_{n-1} = \lambda^2 P_{n-2} = \ldots = \lambda^n P_0$

As $P_0 = p_0 - \dfrac{1}{2} = \dfrac{1}{2}$, this gives $p_n = \dfrac{1}{2} + \dfrac{1}{2}(q - p)^n$

Generating Function : $\qquad G(t) = \displaystyle\sum_{n=0}^{\infty} p_n t^n = \frac{1}{2}\sum_{0}^{\infty} [1 + (q - p)^n]\, t^n.$

Problems with Solutions Provided at the End of the Text

1*. If X_i, $i = 1, 2, 3, 4, 5$ are independent *point* binomial variates each with parameter $p = \frac{1}{2}$, find (a) $P(\overline{X} = 3/5)$, (b) $P(\overline{X} > 3/5)$, (c) $(\overline{X} < 3/5)$.

2*. Let $X \sim$ bin (n, P) and P has the distribution $dF_P(t) = f_P(t)\,dt$, $0 \le t \le 1$. Show that if $E(P^r) = \mu_r'$, then $E(X^{(r)}) = n^{(r)}\,\mu_r'$.

3*. Each of n independent variates X_i, $i = 1, 2, .., n$ takes the value 1 with probability p and the value 0 with prob. $q = 1 - p$. (X_i are indicaters). Find the probability that $X = X_1 + X_2 + \ldots + X_n = x$, when $x = 0, 1, 2, \ldots, n$.

4*. Let $X \sim$ bin (n, p), $Y \sim$ bin (n, q) be indep. $p + q = 1$. Show that $X + Y$ is distributed like $X_1 + X_2 + \ldots + X_n$, where X_i are i.i.d. variates with the Dist. :

$$P(X_i = 0) = pq, \; P(X_i = 1) = p^2 + q^2, \; P(X_i = 2) = pq.$$

5*. Find the p.g.f. of **zero-truncated** bin (n, p). Also obtain variance of such a r.v.

12-70. *Relative (or Pseudo) Binomial Variate*

Let Y be bin (n, p). Then $X = Y/n$, is the ratio of the number of occurrences X to the total number of trials n. The variate Y is called *Relative binomial variate.* Obviously

$$P(X = x) = P(Y = nx) = {}^nC_{nx}\, q^{n-nx} p^{nx}\,; x = 0, \frac{1}{n}, \frac{2}{n}, \ldots, 1; P(X = 0) = 0, \text{elsewhere.}$$

To domain of $f(x\,; n, p)$ is the set of $(n + 1)$ points lying in the closed interval $[0, 1]$. It is also transparent that

$$E(X) = E\left(\frac{Y}{n}\right) = \frac{E(Y)}{n} = \frac{np}{n} = p;\; \text{Var}(X) = \text{Var}\left(\frac{Y}{n}\right) = \frac{\text{Var } Y}{n^2} = \frac{npq}{n^2} = \frac{pq}{n}.$$

Note. X does not have binomial p.m.f. since it is not an integer-valued variate.

Example : Let X be Relative bin (5, 1/4). Find

(a) $P(X = 1/4)$, (b) $P(X = 1/5)$, (c) $P(X = 2/5)$, (d) $P(X > 1/2)$.

Solution. Here $n = 5$, $p = 1/4$, $q = 3/4$ so that $x = \{0, 1/5, 2/5, 3/5, 4/5, 1\}$.

(a) $P(X = 1/4) = 0$, since $1/4$ does not belong to the domain of D.

(b) $$P\left(X = \frac{1}{5}\right) = \binom{5}{1}\left(\frac{3}{4}\right)^4\left(\frac{1}{4}\right)^1 = \frac{405}{1024}.$$

(c) $$P\left(X = \frac{2}{5}\right) = \binom{5}{2}\left(\frac{3}{4}\right)^2\left(\frac{1}{4}\right)^3 = \frac{135}{512}.$$

(d) $$P(X > 1/2) = P(X = 3/5) + P(X = 4/5) + P(X = 1)$$

$$= \binom{5}{3}\left(\frac{3}{4}\right)^2\left(\frac{1}{4}\right)^3 + \binom{5}{4}\left(\frac{3}{4}\right)\left(\frac{1}{4}\right)^4 + \binom{5}{5}\left(\frac{1}{4}\right)^5 = \frac{(90 + 15 + 1)}{1024} = \frac{53}{512}.$$

Exercise 12(c)

1. Bring out fallacy, if any in the statements :
   (a) The mean of a binomial distribution is 5 and the standard deviation is 4.
   (b) The mean of a binomial distribution is 3 and variance is 4.
2. If $X \sim$ bin (n, p) and $Y = X^2$, find Corr (X, Y).
3. Identify the distribution with the m.g.f. $(0.3 + 0.7\, e^t)^4$.
4. (a) If X is bin (n, p) with $E(X) = 5$ and $\sigma_X^2 = 4$, find n and p.

   (b) If X is bin, with index $n = 10$, find the parameter p if $P(X = 2) = \frac{1}{2} P(X = 3)$.
5. (a) If X is bin (25, 0.2), show that $P(X < \mu_x - 2\sigma_x) = (0.8)^{25}$.
   (b) If $X \sim$ bin (100, 0.1), show that $P(X \le \mu - 3\sigma) = (10.9)\,(9/10)^{99}$.
6. If X is bin (5, 1/2), verify that $P(X \le \mu_x) = 1/2$.
7. Determine the binomial distribution for which the mean is 3 and variance 2 and find its mode.
8. Show that the binomial distribution for which (mean + 2 variance) = 4, mean + variance = 3 is bin (4, 1/2).
9. Let $X \sim$ bin (n, p). Show that the mean μ and variance σ^2 can never satisfy the relations $\mu\sigma = 4$, $\sigma/\mu = 2$.
10. Let $X \sim$ bin (n, p). Find n and p if $\mu\, \sigma = 4$ and $\mu/\sigma = 2$. Hence find β_1, β_2 and mode.
11. The mean and variance of a binomial variate X are 4 and 4/3 respectively. Show that $\beta_1 = \cancel{1/2}$, $\beta_2 = 11/4$, $x_0 = 4$, $\delta(\mu) = 640/729$ and $P(X \ge 1) = 728/729$.
    1/12
12. The mean and variance of a binomial variate X are 16 and 8 respectively. Find β_1, β_2, mode, mean deviation $\delta(\mu)$, skewness and kurtosis of the distribution. Find also (i) $P(X = 1)$, (ii) $P(X = 2)$, (iii) $P(X \ge 2)$ and (iv) $P(X \ge 2 \mid X > 0)$.
13. If the m.g.f. of a variate X is $\left(\frac{1}{2} + \frac{2}{3} e^t\right)^5$, show that $P(X = 2) = 40/243$.
14. If X and Y are independent bin (5, 1/2) and bin (7, 1/2) respectively, show that
    $$P(X + Y = 3) = 55/(2)^{10}.$$
15. If X is bin (4, 1/2), show that $P\{\,|X - 2| \ge 1\} = 5/8$, and compare it with Chebyshev's inequality estimate.
16. Use Chebyshev's inequality to determine how many times a fair coin must be tossed in order that the probability will be at least 0.90 that the ratio of the observed number of heads to the number of tosses will lie between 0.4 and 0.6.
17. If X is bin (40, 1/2), use Chebyshev's inequality or otherwise to
    (a) Find k such that $P(\,|X - 20| > 10\,k) \le 0.25$. $[k^2 = 4/10]$
    (b) Show that a lower bound for $P\{\,|X - 20| \le 5\}$ is 0.6.
    (c) Find the upper bound for $P\{\,|X - 10| > 10\}$
    [For (c), use $P\{\,|X - b| > c\} \le E(X - b)^2/c^2$. Here $\mu = 20$, $\sigma^2 = 10$, $E(X - 10)^2 = 110$, bound = 110/100 > 1, a useless bound]
18. Assume that you have a fair die. How many times should you roll it (what is the least number of tosses) so that with a probability of at least 85% the frequency of 6's will differ from 1/6 by less than 0.03 ?

19. You have a die but you do not know whether it is loaded or not. How many times should you roll it so that with a probability of at least 85% the frequency of the 6's will differ from the actual (but unknown) probability of getting a 6 by less than 0.05 ?

20. There is accommodation for 5 persons in an air conditional compartment of a railway carriage. The number of accepted seats in this carriage follows a binomial dist. with $p = 0.06$. For each seat the railway have a net income of Rs. 50/-, and for each vacant seat there is a net loss of Rs. 15/-. Find the average net gain for the railway by running this carriage.

21. A quality control engineer takes daily samples of 10 electronic components and checks them for inspections. On 200 consecutive working days, he obtained 112 samples with zero defective, 76 samples with one defective and 12 samples with two defectives. If these samples can be looked upon as samples from binomial population, obtain the theoretical or expected number of samples with zero defective, one defective, and 2 or more defective.

22. If the probability that a child is son is p, where $0 < p < 1$, find the expected number of sons in a family with n children, given that there is at least one son.

23. Determine the expected number of boys in a family with 8 children, assuming the sex distribution to be equally probable. What is the probability that the expected number of boys does occur ?

24. The probability that a defective soldered connection is made on any given connection is $1/10^4$. Find the expected number and S.D. of defective joints in a system with 5×10^4 soldered connections. Also find the probability that there are no defects in the system.

25. If X is bin (n, p) show that X is symmetrically distributed about c iff $p = 1/2$, $c = n/2$.

26. X_1, X_2, X_3 are independent Bernoulli variates each having the parameter $p = 1/4$. Show that

(a) $P(\Sigma X_i = 2) = 9/64$ (b) $P(\overline{X} < 1) = 63/64$.

27. X_i, $i = 1, 2, 3, 4$ are independent Bernoulli variates each having the parameter $p = 1/2$. Show that (a) $P(\overline{X} = 3/4) = 1/4$. (b) $P(\overline{X} > 3/4) = 1/16$.

28. Determine the binomial distribution when

(a) $\beta_1 = 1/36, \quad \beta_2 = 35/12$ (b) $\beta_1 = 1/15, \quad \beta_2 = 89/30$.

[**Ans.** (a) bin (18, 1/3) or bin (18, 2/3); (b) bin (20, 1/4) or bin (20, 3/4)]

29. If X is bin (n, p) show that

(a) $P\{X \text{ is even}\} = (1/2)[1 + (q - p)^n]$. (b) $k_5 = npq\,(q - p)(1 - 12\,pq)$.

30. If X is bin (n, p) prove that Recursion formula

$$f_n(x) = \frac{p}{q}\,\frac{n - x + 1}{n}\, f_n(x - 1), \ (x \ge 1, p \ne q)$$

If $p = q$, this reduces to the form $f_n(x) + f_n(x - 1) = 2f_{n-1}(x)$.

Using the factorial moments, or otherwise, prove that $\gamma_2 = 0 \Rightarrow p$ or $q = \frac{1}{2}(1 + 1/\sqrt{3})$.

31. If X and Y are independent bin (n_1, p_1) and bin (n_2, p_2) ; find $P(X + Y = k)$. What is the distribution of $X + Y$ if $p_1 = p_2$?

32. If X and Y are two i.i.d. binomial variates, obtain $P(|X\ Y| = r)$, r : positive integer.

33. If X is bin (n, p), show that $P(X \le x) = \left(\int_{p/q}^{\infty} \frac{y^x}{(1 + y)^{n+1}}\, dy\right) \Big/ \left(\int_0^{\infty} \frac{y^x}{(1 + y)^{n+1}}\, dy\right)$.

34. If a random variable takes values 0 and 1 only; show that all its moments about 0 are equal.

35. Let $X \sim$ bin $(1, p_1)$ and $Y \sim$ bin $(1, p_2)$ be independent. Find the density of $Z = XY$. Verify your answer by calculating the mean and variance of Z by two ways.

36. Let X be bin (n, p) and r be a non-negative integer. If the rth moment about the origin is denoted by $\mu'_r = \mathsf{E}(X^r)$, prove that

$$\mu'_{r+1} = np\mu'_r + pq\, d\mu'_r/dp.$$

12-80. Miscellaneous Worked-out Problems

***Example* 1.** If X and Y are independent Bernoulli variates with parameters 1/2, show that $X + Y$ and $|X - Y|$ are dependent though uncorrelated.

Solution. Here $P(X = 1) = 1/2 = P(X = 0)$; $P(Y = 1) = 1/2 = P(Y = 0)$. Let $S = X + Y$ and $D = |X - Y|$. The distributions of S and D are easily seen to be

$$P(S = 0) = 1/4,\ P(S = 1) = 1/2,\ P(S = 2) = 1/4,\ P(D = 0) = P(D = 1) = 1/2.$$

$$P(S = 0, D = 0) = P(X = 0, Y = 0) = P(X = 0)\, P(Y = 0) = 1/4.$$

$$P(S = 0) = P(X = 0, Y = 0) = P(X = 0)\, P(Y = 0) = 1/4.$$

$$P(D = 0) = P(X = 1, Y = 1) + P(X = 0, Y = 0) = 1/4 + 1/4 = 1/2.$$

$$\therefore \quad P(S = 0, D = 0) \neq P(S = 0)\, P(D = 0) \quad \Rightarrow \quad S \text{ and } D \text{ are dependent.}$$

Now, $\mathsf{E}(S) = 0.1/4 = 1\,.\,1/2 + 2\,.\,1/4 + 1, \quad \mathsf{E}(D) = 0\,.\,1/2 + 1\,.\,1/2 = 1/2$

$\therefore$ Cov $(S, D) = \mathsf{E}(SD) - \mathsf{E}(S)\,\mathsf{E}(D) = 1/2 - 1/2 = 0 \quad \Rightarrow \quad S$ and D are uncorrelated.

***Example* 2.** Show that $\mathsf{E}(X + a)^{-1} = \int_0^1 t^{a-1} G(t)\, dt,\ a > 0$. ...(1)

Find it when $X \sim$ bin (n, p) and a = 1. Also evaluate bias of $Y = (n + 1)/(X + 1)$, as an estimate of $(1/p)$.

Solution. Formally : $G(t) = \mathsf{E}(t^X)$, so

$$\int_0^1 t^{a-1} G(t)\, dt = \int_0^1 t^{a-1} \Big[\sum_x t^x f(x) \Big] dt = \sum_x f(x) \int_0^1 t^{a+x-1}\, dt = \sum_x \frac{f(x)}{(a+x)} = \mathsf{E}\,(X + a)^1 .$$

For bin (n, p) $G(t) = (q + pt)^n$ and with $a = 1$, Eq. (1) provides

$$\mathsf{E}\left(\frac{1}{X+1}\right) = \int_0^1 (q + pt)^n\, dt = \frac{1}{(n+1)\,p} \left\{ [q + pt]^{n+1} \right\}_0^1 = \frac{1 - q^{n+1}}{(n+1)\,p}. \qquad ...(2)$$

Bias $= \mathsf{E}(Y) - (1/p) = (n + 1)\,\mathsf{E}(X + 1)^{-1} - (1/p) = -(q^{n+1}/p)$, [by (2)].

Obviously, bias $\to 0$ as $n \to \infty$, since $q^n \to 0$ as $n \to \infty$, $0 < q < 1$.

***Example* 3.** Let $X_1, X_2, ..., X_n$ be i.i.d. bin (1, 1/2) variates. Find $P(Y = 0, Z = 1)$ where $Y = \min\{X_1, ..., X_n\}$ and $Z = \max\{X_1, ..., X_n\}$.

Solution. Let $A = \{Y = 0\}$ and $B = \{Z = 1\}$; then $\overline{A} = \{Y = 1\}$ and $\overline{B} = \{Z = 0\}$, $(\overline{A} \cap \overline{B} = \varnothing)$

$$P(\overline{A}) = P\{\min(X_1, ..., X_n) = 1\} = P\{X_1 = 1, ..., X_n = 1\} = [P(X_1 = 1)]^n = (1/2)^n.$$

$$P(\overline{B}) = P\{\max(X_1, ..., X_n) = 0\} = P\{X_1 = 0, ..., X_n = 0\} = [P(X_1 = 0)]^n = (1/2)^n.$$

$$P(AB) = 1 - P(\overline{A \cap B}) = 1 - P(\bar{A} \cup \bar{B}) = 1 - [P(\bar{A}) + P(\bar{B})]$$

$$= 1 - [(1/2)^n + (1/2)^n] = 1 - (1/2)^{n-1}.$$

***Example* 4. Simple Random Walk.** A drunk performs a random walk over positions 0, ±1, ±2, ... as follows : He starts at 0. He takes successive one-unit steps, going to the right with probability p and to the left with probability $q = 1 - p$. His steps are independent. Let X denote his position after n steps. Find the distribution of $Y = (X + n)/2$ and Var (X). Also find $P(X = 0)$.

Solution. If X_i is the ith step (movement), then by hypothesis $P(X_i = 1) = p$, and $P(X_i = -1) = q$. Let $X = X_1 + X_2 + \dots + X_n$; then $-n \le x \le n$. Now

$$G(t : X_i) = E(t^{X_i}) = pt + qt^{-1}.$$

$$G[t : (X + n)/2] = t^{n/2} G(t^{1/2} : X) = t^{n/2}\{G(\sqrt{t} : X_i\}^n = t^{n/2}(p\sqrt{t} + q/\sqrt{t})^n = (q + pt)^n.$$

This shows that $Y = (X + n/2) \sim$ bin (n, p) ; the distribution sought. Now

$$E[X + n)/2] = np, \text{ Var } [(X + n)/2] = npq \quad \Rightarrow \quad E(X) = n(p - q) \text{ ; Var } (X) = 4npq.$$

Now $\quad p_0 = P(X = 0) = P\left(Y = \frac{n}{2}\right) = \binom{n}{\frac{1}{2}n} p^{n/2} q^{n/2}.$

Thus, $\quad p_0 = \dfrac{n(n-1)\dots(\frac{1}{2}n + 1)}{(n/2)!}(pq)^{n/2}$, if n is even; $p_0 = 0$ if n is odd.

This is the probability that the drunk shall be at the origin after n steps.

***Example* 5.** If $X \sim$ bin (n, p) and $Y \sim B_1$ $(k, n - k + 1)$, then $F_Y(p) = 1 - F_X(k - 1)$. That is

$$P\{X \ge k\} = [B(k, n-k+1)]^{-1} \int_0^p u^{k-1}(1-u)^{n-k}\, du \qquad \dots(1)$$

Deduce : $P(X \ge k) \le \binom{n}{k} p^k.$

Solution. We differentiate w.r.t. 'p' the survival function $P \equiv P(X \ge k) = \sum_{r=k}^{n} \binom{n}{r} q^{n-r} p^r$

$$\frac{dP}{dp} = \sum_{r=k}^{n}\left\{r\binom{n}{r}q^{n-r}p^{r-1} - (n-r)\binom{n}{r}q^{n-r-1}p^r\right\}, \qquad \left[(n-r)\binom{n}{r} = n\binom{n-1}{r}\right]$$

$$= \sum_{r=k}^{n}\left\{n\binom{n-1}{r-1}q^{n-r}p^{r-1} - n\binom{n-1}{r}q^{n-r-1}p^r\right\}$$

$$= n\sum_{r=k}^{n}\{\phi(r-1) - \phi(r)\}. \qquad \left[\phi(r) = \binom{n-1}{r}q^{n-r-1}p^r,\ \phi(n) = 0\right]$$

$$= n[\phi(k-1) - \phi(n)] = n\phi(k-1) = n\binom{n-1}{k-1}q^{n-k}p^{k-1} = n\binom{n-1}{k-1}p^{k-1}(1-p)^{n-k}.$$

Integrating this result on $[0, p]$ and adjusting dummy, we get

$$P = n\binom{n-1}{k-1}\int_0^p u^{k-1}(1-u)^{n-k}\, du. \qquad \dots(2)$$

Note : $n\binom{n-1}{k-1} = \dfrac{n!}{(k-1)!(n-k)!} = \dfrac{\Gamma(n+1)}{\Gamma(k)\Gamma(n-k+1)} = \dfrac{1}{B(k, n-k+1)}$ ... (3)

So, using (3) into Eq. (2) yields the stated result (1) : viz. $P\{Y \le p\} = P\{X \ge k\}$.

Deduction : $u > 0 \Rightarrow 1 - u < 1 \Rightarrow (1-u)^{n-k} < 1$.

$$\therefore \qquad P \le n\binom{n-1}{k-1}\int_0^p u^{k-1}\,du = \binom{n}{k}p^k.$$

Example 6. Let $X \sim$ bin (n, p). Show that

$$P(X \le k) = (n-k)\binom{n}{k}\int_0^q t^{n-k-1}(1-t)^k\,dt = 1 - \frac{1}{B(k+1, n-k)}\int_0^p t^k(1-t)^{n-k-1}\,dt.$$

Deduce : $$P\{X \ge k\} = \frac{1}{B(k, n-k+1)}\int_0^p t^{k-1}(1-t)^{n-k}\,dt.$$

Solution. Observe : $(n-k)\binom{n}{k} = \dfrac{n!}{k!(n-k-1)!} = \dfrac{\Gamma(n+1)}{\Gamma(k+1)\Gamma(n-k)} = \dfrac{1}{B(k+1, n-k)}$. ...(1)

(i) Now $\qquad F(k) = P(X \le k) = \sum\limits_{r=0}^{k}\binom{n}{r}q^{n-r}p^r$ ...(2)

Differentiating (2) w.r.to q we obtain

$$\frac{dF}{dq} = \sum_{r=0}^{k}\binom{n}{r}\left[(n-r)q^{n-r-1}p^r - rq^{n-r}p^{r-1}\right]$$

$$= \sum_{r=0}^{k}\left[n\binom{n-1}{r}q^{n-r-1}p^r - n\binom{n-1}{r-1}q^{n-r}p^{r-1}\right], \quad \left\{(n-r)\binom{n}{r} = n\binom{n-1}{r}\right\}$$

$$= n\sum_{r=0}^{k}\{\phi(r) - \phi(r-1)\}, \qquad \left[\phi(r) = \binom{n-1}{r}q^{n-1-1}p^r\right]$$

$$= n\phi(k) \qquad [\phi(-1) \equiv 0] \text{ Telescopic Series.}$$

Integrating both sides w.r.t. q between limits $[a, q]$, we obtain

$$[F(k)] = n\binom{n-1}{k}\int_0^q t^{n-k-1}p^k\,dq = (n-k)\binom{n}{k}\int_0^q t^{n-k-1}(1-t)^k\,dt \qquad \text{...(3)}$$

where the dummy variable q is replaced by dummy variable t and p by $(1-t)$.

$$\because \int_0^q t^{n-k-1}(1-t)^k\,dt = \int_0^1 t^{n-k-1}(1-t)^k\,dt - \int_q^1 t^{n-k-1}(1-t)^k\,dt$$

$$= B(k+1, n-k) - \int_0^p z^k(1-z)^{n-k-1}\,dz, \quad [z = 1-t] \qquad \text{...(4)}$$

From (3) and (4) using (1), we obtain

$$F(k) = 1 - \frac{1}{B(k+1,\ n-k)}\int_0^p z^k(1-z)^{n-k-1}\,dz. \qquad \text{...(5)}$$

(ii) $\qquad P\{X \ge k) = 1 - P\{X < k\} = 1 - P\{X \le k-1\}$

$$= [B(k, n-k+1]^{-1}\int_0^p z^{k-1}(1-z)^{n-k}\,dz. \qquad \text{[by (5)]} \qquad \text{...(6)}$$

***Example* 7.** Let $X_1 \sim$ bin (n, p_1) and $X_2 \sim$ bin (n, p_2). If $p_1 < p_2$, show that
$$P\{X_1 \le k\} \ge P\{X_2 \le k\}, \text{ for } k = 0, 1, ..., n.$$

Solution. The relation between binomial and beta distribution is

$$P\{X \ge r\} = \frac{1}{B(r, n-r+1)} \int_0^p x^{r-1}(1-x)^{n-r}\,dx. \qquad ...(1)$$

$$\therefore \quad P(X \le k) = 1 - P(X > k) = 1 - P(X \ge k+1) = 1 - [B(k+1, n-k]^{-1} \int_0^p x^k (1-x)^{n-k-1}\,dx$$

or $\int_0^p x^k (1-)^{n-k-1}\,dx = \lambda[1 - P(X \le k)] \quad [\lambda = B(k+1), n-k)].[r = k+1] \qquad ...(2)$

If $p_2 > p_1$, then $\int_0^{p_2} f(x)\,dx > \int_0^{p_1} f(x)\,dx$. $[f(x) = x^k(1-x)^{n-k-1}]$ whence

$$\lambda[1 - P(X_2 \le k)] \ge \lambda[1 - P(X_1 \le k)] \Rightarrow P(X_1 \le k) \ge P(X_2 \le k).\ F_1(k) \ge F_2(k)$$

[The smaller the p_1, the more the binomial distribution is shifted to the left].

***Example* 8.** Let $X \sim$ bin (n, p). Show that

$$P(X \le k) = \lambda \int_{p/q}^{\infty} \frac{y^k dy}{(1+y)^{n+1}}, \text{ where } \frac{1}{\lambda} = \int_0^{\infty} \frac{y^k\,dy}{(1+y)^{n+1}} = B(k+1, n-k).$$

Solution. Since $P(X \le k) = \sum_{x=0}^{k} \binom{n}{x} q^{n-x} p^x$,

$$\therefore \quad \frac{d}{dq} P(X \le k) = \sum_{x=0}^{k} \left[\binom{n}{x}(n-x)\,q^{n-x-1}\,p^x - x\binom{n}{x} q^{n-x} p^{x-1}\right]$$

$$= \sum_{x=0}^{k} \left[n\binom{n-1}{x} q^{n-x-1}\,p^x - n\binom{n-1}{x-1} q^{n-x} p^{x-1}\right]$$

$$= \sum_{x=0}^{k} [\phi(x) - \phi(x-1)] = \phi(k).\left\{\phi(x) = n\binom{n-1}{x} q^{n-1-x} p^x, \phi(-1) = 0\right\} \qquad ...(1)$$

$$\frac{1}{\lambda} = \int_0^{\infty} \frac{y^{(k+1)-1}}{(1+y)^{n+1}}\,dy = B(k+1\,n-k) = \frac{k!(n-k-1)!}{n!} = \left\{n\binom{n-1}{k}\right\}^{-1}$$

$$\therefore \quad \frac{d}{dq}\lambda \int_{p/q}^{\infty} \frac{y^k dy}{(1+y)^{n+1}} = \frac{\lambda(p/q)^k}{[1+(p/q)]^{n+1}} \cdot \frac{1}{q^2} = \lambda p^k q^{n-k-1} = \phi(k) \qquad \text{[by (1)]}$$

Thus $\frac{d}{dq}\left[P(X \le k) - \lambda \int_{p/q}^{\infty} \frac{y^k dy}{(1+y)^{n+1}}\right] = 0$. We integrate it to get

$$P(X \le k) = \lambda \int_{p/q}^{\infty} \frac{y^k\,dy}{(1+y)^{n+1}},$$

the constant of integration vanishing when $q = 0$.

Remark. $P\{\text{bin } (n, p) \le k\} = P\{B_{II}(k+1, n-k) \ge p/q\}$.

Example 9. Independent binomial variates X_i $(i = 1, 2, ..., k)$ have parameters n and p_i. Find the mean and variance of $X = X_1 + ... + X_k$ and compare these with the mean and variance of a binomial variate Y with parameters nk and $p = \Sigma(p_i/k)$.

Solution. Since $X_i \sim$ bin (n, p_j), we know that $E(X_j) = np_j$, Var $(X_j) = np_j q_j$ so that

$$E(X) = E(X_1) + E(X_2) + ... + E(X_k) = n(p_1 + p_2 + ... + p_k) = nkp.$$

$$\text{Var}(X) = \text{Var}(X_1) + \text{Var}(X_2) + ... + \text{Var}(X_k) = n[p_1 q_1 + p_2 q_2 + ... + p_k q_k]$$

$$= n\,\Sigma p_i\, q_i = n\,\Sigma p_i\,(1 - p_i) = n(\Sigma p_i) - n\,(\Sigma p_i^2) = nkp - n\Sigma p_i^2$$

Now $E(Y) = nkp$, $\quad$ Var $(Y) = nkpq$

$$\therefore \quad \frac{\text{Var}(X)}{\text{Var}(Y)} = \frac{npk - n\,\Sigma p_i^2}{knpq} = \frac{1 - [(\Sigma p_i^2)/pk]}{q}$$

From $\displaystyle\sum_{i=1}^{k}(p_i - p)^2 = \sum_{i=1}^{k} p_i^2 - kp^2 \geq 0, \Rightarrow \frac{\Sigma p_i^2}{kp} \geq p \Rightarrow 1 - \left(\frac{\Sigma p_i^2}{kp}\right) \leq 1 - p = q$

$\therefore$ Var (X)/Var $(Y) \leq 1$, i.e. Var $(X) \leq$ Var (Y).

This result is somewhat unusual.

Problems with Solutions Provided at the End of the Text

1*. Show that if two symmetrical binomial distributions $(p = q = 1/2)$ of degree n (and of the same number of observations) are so superimposed that the rth term of one coincides with $(r + 1)$th term of the other, the distribution formed by adding superimposed terms is a symmetrical binomial distribution of degree $(n + 1)$.

2*. Random variable X is bin (n, p) but the parameter p takes one of the two values p_1 and p_2 with respective probabilities θ and $1 - \theta$. Find the probability that $X = n$ given $X > 0$.

3*. The failure rate of a new process is estimate at 75%. How many times should the process be run to given an 80% chance of at least two successes ?

4*. For one half of n events, the chance of success is p, and the chance of failure is q, whilst for the other half the chance of success is q, and the chance of failure is p. Show that the S.D. of the number of successes is the same as if the chance of success were p in all the cases i.e. $\sqrt{npq}$, but that the mean of the number of successes is $n/2$ and not np.

5*. Show that $\mu_k' = (p\partial / \partial p)^k\,(q + p)^n$.

6*. Let (X, Y) have joint p.d.f.

$$P(X = x, Y = y) = \left(\frac{(1-\theta)\,\theta'}{\theta + \theta'}\right)^{(1-x)(1-y)} \cdot \left(\frac{\theta\theta'}{\theta + \theta'}\right)^{x(1-y)+y(1-x)} \left(\frac{\theta(1-\theta')}{\theta + \theta'}\right)^{xy}$$

where $x = 0, 1$; $y = 0, 1$; $0 \leq \theta, \theta' \leq 1$.

Find $f_X(x)$, $f_Y(y)$ and Corr (X, Y).

Miscellaneous Exercises

1. With the usual notation, find p for a binomial variate X when $n = 6$ and if $P(X = 4) = P(X = 2)$.
2. X is a binomial variate with mean 6 and variance 4. Find n and p. Hence evaluate (i) $P(X > 15)$, (ii) $P(5 < X < 10)$, (iii) $P(X \le 8)$, (iv) $P(X \ge 6)$, (v) $P(2 \le X \le 12)$, (vi) $P(3 < X \le 16)$ and $P(X = 4 \mid X > 0)$.
3. In a binomial distribution consisting of 5 independent trials, probabilities of 1 and 2 successes are 0.4096 and 0.2048 respectively. Find the parameter p of the distribution.
4. For a Binomial distribution the mean is 6 and the standard deviation is $\sqrt{2}$. Write out all the terms of the distribution.
5. Let X be bin $(2, p)$ and Y be bin $(4, p)$. If $P(X \ge 1) = 5/9$, find $P(Y \ge 1)$.
6. If the m.g.f. of a variate X is $\left(\frac{2}{3} + \frac{1}{2} e^t\right)^9$ find $P(X = 5 \text{ or } 6)$. Also show that

$$P(\mu - 2\sigma < X < \mu + 2\sigma) = \sum_{x=1}^{5} \binom{9}{x} \left(\frac{1}{3}\right)^x \left(\frac{2}{3}\right)^{9-x}.$$

7. Let X be the number of successes throughout n independent repetitions of a random experiment having prob. of success $p = 1/4$. Show that the smallest value of n so that $P(X \ge 1) \ge 0.70$, is 4.
8. Let the independent variates X_1 and X_2 have binomial distributions with parameters $n_1 = 3$, $p_1 = 2/3$ and $n_2 = 4$, $p_2 = 1/2$ respectively. Compute $P(X_1 = X_2)$.
9. Let $X \sim$ bin (n, p). If $F(x, n, p)$ is c.d.f. of X then show that $F(\lambda, n, p) > 1/2$ if $p \le 1/2$ and $\lambda = np$ is an integer.
10. Let $X \sim$ bin (n, p). What is the distribution of $n - X$?
11. Let $X_1 \sim$ bin (n_1, p) and $X_2 \sim$ bin (n_2, q) be independent. What is the distribution of $X_1 - X_2 + n_2$?
12. If $X \sim$ bin (n, p), show that $P(X \le 2) = P(X \ge n - 2)$ iff $p = 1/2$.
13. In a sequence of independent trials, the probability of an event A varies cyclically through values $p, p + \delta, p - \delta, p + \delta, \ldots$, where $0 < p - \delta < p < p + \delta < 1$.
    (a) Find the probability of observing r A's (exactly) in 4 successive trials.
    (b) Find the mean and variance of the distribution of r, and compare with the corresponding values for a binomial distribution with probability p.
14. A machine usually makes items of which 4% are defectives. Every hour, a sample of size 10 is drawn for inspection. If this sample contains no defectives, machine is not stopped. Show that the chance that the machine is not stopped when it has started producing 10% defectives is $(0.9)^{10}$.
15. If $X \sim$ bin (N, p) and if N is allowed to vary, then show that $E(X) = p\,E(N)$ and Var $(X) = p\,E(N) + p_2\,[\text{Var}(N) - E(N)]$.
16. Let $X_1, X_2, \ldots, X_n$ be i.i.d. bin $(1, p)$ variates. Find the p.d.f. of $Y = \sqrt{X_1^2 + X_2^2 + \ldots + X_n^2}$.
17. The prob. p_n that n persons visit a supermarket in one day is $p_n = qp^n$, $n = 0, 1, 2, \ldots$ Two out of three persons, on average, buy an item D. The prob. that D is defective is 1/4. Show that P {a person buys a defective D} = 1/2. If k non-defective items are sold, show that the conditional prob. r_n that n persons visited the shop is

$$r_n = {}^nC_k\, p^{n-k} (2 - p)^{k+1} / 2^{n+1}.$$

18. If $X \sim$ bin (n, p), prove that : $E[(1 + X)^{-1}] = [1 - q^{(n+1)}] / (n + 1)p$.

$$E[(1+X)^{-1}(2+X)^{-1}] = [1-q^{n+2}-(n+2)pq^{n+1}]/(n+1)(n+2)p^2.$$

19. Let the zero truncated bin (n, p) be defined by

$$P(X=x) = {}^nC_x\, p^x q^{n-x}(1-q^n)^{-1},\ x = 1, 2, \ldots, n; \qquad P(X=0) = 0,$$

Show that the negative moments $M_{-k} = E(X^{-k}),\ k = 1, 2, \ldots,\ M_0 = 1,$ satisfy the equations

$$\frac{dM_{-k}}{dq} = \frac{nM_{-k}}{q(1-q^n)} - \frac{M_{1-k}}{q(1-q)}.$$

Hence obtain the lower moments.

20. Let X be a binomial variate and let $B(k; n, p) = \sum_{x=0}^{k} b(x; n, p) = \sum_{x=0}^{k} \binom{n}{x} p^x q^{n-x}$.

Show that the following (tail-probability results) hold

(a) $1 - B(k-1, n, p) \le \dfrac{nb(k; n, p)}{k-np},\ k > np+1$ (b) $B(k, n, p) \le \dfrac{n}{np-k}\, b(k; n, p),\ k < np$

(c) $1 - B(k; n, p) = n\binom{n-1}{k}\int_0^p t^k (1-t)^{n-k-1}\, dt.$

21. Ten students are chosen at random from a college at which 20 percent of the students do not reside in the college compound. Find the probability that exactly 8 of the students do not reside in the college compound. What is the expected number of non-resident students in the sample ?

22. The probability of team A winning any game is 1/2. What is the probability that A wins more than half of its games if A plays n times, when $n = 2m + 1,\ n = 2m$?

23. If a fair coin is tossed at random five independent times, find the conditional probability of 5 heads relative to hypothesis that there are at least four heads.

24. The probability that any of n identical units takes part in an experiment is p, $(p\ 1/n)$. If a given unit participates in the experiment exactly k times, the result of these experiments is considered attained. Find the probability of attaining the desired result in m experiments.

Find also the prob. of attaining the desired result in $(2k - 1)$ experiment, if the experiments are discontinued when the result has been attained.

25. Two independent sequence of independent trials are performed. The first sequence consists of n_1 trials and the second consists of n_2 trials. The probability of a success on any trial in the first sequence is p_1 and on any trial in the second sequence is p_2. What is the probability of a total of k successes in the combined trial ?

26. If in a Bernoulli sequence of n trials, it is known that there are exactly k successes, what is the conditional probability of a success on jth trial ?

27. Two dice are thrown n times. Let X denote the number of throws in which the number on the first die exceeds the number on the second die. Show that X is bin (n, p) where $p = 5/12$.

28. Suppose n numbered balls $(1, 2, \ldots, n)$ are randomly distributed among n cells, also numbered $1, 2, \ldots, n$. If X be the number of matches, show that X is bin $(n, 1/n)$.

29. Starting at the origin a particle takes steps of L-units length, to the right with probability p and to the left with probability $q = 1 - p$. Assuming independent trials, find the probability after N steps and also variance of the distance.

30. Assume that on the average one telephone number out of fifteen called between 2 P.M. and 3 P.M. on Sunday is busy. What is the probability that if six randomly selected telephone numbers are called

(a) not more than three, (b) at least three of them will be busy ?

31. The probability that a unit must undergo repairs after k accidents is given by the formula $f(k) = 1 - (1 - 1/\theta)^k$, where θ is the average number of accidents before the unit is submitted for repairs. Prove that the probability that after n cycles the unit will need repairs is given by $P_n = 1 - (1 - p/\theta)^n$, where p is the probability that an accident will occur during one cycle.

32. A point A must be connected with 10 telephone subscribers at a point B. Each subscriber keeps the line busy 12 minutes per hour. The calls from any two subscribers are independent. Find the minimal number of channels necessary, so that all the subscribers will be served at any instant with probability 0.99.

34. Independent binomial variates X_i $(i = 1, 2, ..., k)$ have parameters n and p^i. Find the mean and Var. of $\overline{X} = \Sigma(X_i / k)$ and compare these with the mean and variance of the bin $[n, \Sigma(p^i/k)]$.

The opportunity God sends does not wake him up who is asleep.

We are all born mad, some ever remain so.

Poisson Distribution

13

13-10. *Random Processes, Counting Processes, Poisson Distribution*

1. *Definition* : A **random process** is a family of variates $\{X(t) : t \in T\}$ defined on a given probability space S, indexed by the parameter $t \in T$, where T is the parameter set of the random process.

Since a r.v. is a function defined on S, it follows that the random process $\{X(t) : t \in t\}$ is really a function of two arguments : precisely $\{X(t, \omega), t \in T, \omega \in S\}$. For a fixed $t (= t_k)$, $X(t_k, \omega) = X_k(\omega)$ is simply variate $X(t_k)$, as $\omega \in S$. And for a fixed $\omega_i \in S$, $X(t, \omega_i) = X_i(t)$ is a one-variable function of time t called a **sample function**. If both arguments are fixed, $X(t_k, \omega_i)$ is simply a real number.

2. *Definition* : A random process $\{X(t), t \geq 0\}$ is said to be a **counting process** if $X(t)$ represents the total number of events which have occured in the time interval $(0, t)$.

The counting process $X(t)$ must satisfy the following requirements :

1. $X(t) \geq 0$ and $X(0) = 0$.
2. $X(t)$ is integer-valued function.
3. $X(t_1) \geq X(t_2)$ if $t_2 > t_1$.
4. $X(t_2) - X(t_1)$ = Number of events that have occured on interval (t_1, t_2).

3. *Definition* : A random variable X is said to be a Poisson variate (or X possesses Poisson distribution) if its probability law $P(X = x) \equiv f(x)$ is specified by

$$f(x) = e^{-\lambda} \lambda^x / x!, \quad x = 0, 1, 2, ..., \infty, \qquad f(x) = 0, \text{ otherwise} \qquad ...(1)$$

The constant $\lambda > 0$ is called the *parameter* of the distribution. Obviously

$$\sum_{x=0}^{\infty} f(x) = e^{-\lambda} \sum_{x=0}^{\infty} \frac{\lambda^x}{x!} = e^{-\lambda} e^{\lambda} = 1.$$

The c.d.f. is $F_X(x) = P\{X \leq x\} = \Sigma f(t), \ t \leq x$.

Note. The parameter $\lambda = E(X) = \text{Var}(X)$ [§ 8-31]

Poisson trials. These are characterised by the following facts :

1. The result of each trial is classified into two categories : say success (S) and failure (F).
2. The outcomes of any trial is independent of the outcomes of other trials.
3. The probability $P(S)$ varies from trial to trial.
4. The series of trials is performed an infinite number of times.

Poisson Counting Process

Any process which obeys the following five assumptions is called Poisson Counting Process :

A_1 : The number of successes during non-overlapping time intervals are independent variates.

A_2 : The number of successes depends on the length of that interval.

A_3 : The probability that a success occurs during the time interval $(t, t+\Delta t), \Delta t \ll 1$ is $\lambda \Delta t$, except for infinitesimals of higher order, with λ = constant. ($\lambda > 0$ is called **rate** or intensity).

A_4 : The probability of obtaining more than one success during time Δt is negligible.

A_5 : Let $f_k(t) = P$ {in time t exactly k successes occurs} : then

$f_0(0) = 1$ (at the start $t = 0$, it is certain that no success has occured),

$f_k(0) = 0$ for $k \geq 0$.

[*Note.* $\lambda \Delta t \to 0$ as $\Delta t \to 0$]

13-11. *Derivation of Poisson pmf through Poisson Counting Process*†

In order to have k successes in time $t + \Delta t$, we must either have k success in time t followed by no success in time Δt (with prob = $1 - \lambda$ Δt), or $k - 1$. Successes in time t followed by a single success in time Δt. We need not bother about having two or more successes in time Δt [by A_4]. Hence

$$f_k(t+\Delta t) = f_k(t)[1-\lambda\Delta t] + f_{k-1}(t)\lambda\Delta t \quad \Rightarrow \quad \frac{f_k(t+\Delta t) - f_k(t)}{\Delta t} = -\lambda f_k(t) + \lambda f_{k-1}(t).$$

Letting $\Delta t \to 0$, we get the differential-difference equation :

$$Df_k(t) + \lambda f_k(t) = \lambda f_{k-1}(t), \qquad k = 0, 1, 2, \ldots (D = d/dt). \qquad \ldots(1)$$

We need solve this equation, subject to the conditions

$$f_{-1}(t) = 0, \ \forall \ t; f_0(0) = 1, \ k > 0; \ f_k(0) = 0, k > 0.$$

Assume the solution of (1) is given by the power series :

$$G(t, \theta) = \sum_{k=0}^{\infty} f_k(t)\theta^k \qquad \ldots(2)$$

Since $\Sigma f_k = 1$, (k = 0, 1, ...), the power series is necessarily convergent for $0 \leq \theta \leq 1$. We multiply (1) by θ^k and sum over k = 0, 1, 2, ... to obtain

$$\sum_{k=0}^{\infty} Df_k(t)\theta^k + \lambda \sum_{k=0}^{\infty} f_k(t)\theta^k = \lambda\theta \sum_{k=0}^{\infty} f_{k-1}(t)\theta^{k-1}$$

i.e. $$\frac{\partial g}{\partial t} + \lambda G = \lambda\theta G \quad \text{or} \quad \frac{1}{G}\frac{\partial G}{\partial t} = \lambda(\theta - 1). \qquad \left[\because \frac{\partial G}{\partial t} = \sum_{k=0}^{\infty} \frac{df_k(t)}{dt}\theta^k, \text{ formally}\right]$$

A simple integration provides : $\ell n\, G(t, \theta) = \lambda(\theta - 1)t + A$(const.)

At $t = 0$, $G(0, \theta) = \sum_{k=0}^{\infty} f_k(0)\theta^k = f_0(0) = 1 \Rightarrow A = \ell n\,(1) = 0$

† The reader may skip over §13-11 on first reading. It is not used subsequently.

$$\therefore \quad G(t, \theta) = e^{-\lambda t} e^{\lambda \theta t} = e^{-\lambda t} \sum_{k=0}^{\infty} \frac{(\lambda t)^k}{k!} \theta^k \qquad \text{...(3)}$$

Comparing (2) with (3) we find that : $f_k(t) = e^{-\lambda t}(\lambda t)^k / k!$, $\quad k = 0, 1, 2, \ldots$ ...(4)

For $t = 1$ (Unit time), this is : $f_k = e^{-\lambda} \lambda^k / k!$ (Poisson density), $k = 0, 1, 2, \ldots$

Note. Equation (4) is called Two-*Parameter Poisson density.*

Problem. Prove that, under certain conditions to be stated by you, the number of telephone calls on a trunk line in a given interval of time, has a Poisson distribution. [Hint : Art 13-11].

Illustration. Let Z_j, $j = 1, 2, \ldots$, be i.i.d. variates with p.m.f. $= P(Z_n = 1) = p$, $P(Z_n = -1) = q$, $(p + q = 1)$. Write

$$X_0 = 0, X_n = Z_1 + Z_2 + \ldots + Z_n, \quad n = 1, 2, 3, \ldots$$

The family $\{X_n, n \geq 0\}$ is a random process, called *simple random walk* $X(n)$ in one dimension (notated : RW). Describe the simple RW, $X(n)$ and construct a typical realization of $X(n)$. Find Var (X_n).

Approach. The simple RW, $X(n)$ is discrete-state random process. The state space is $E = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$ and index parameter. Set $T = \{0, 1, 2, \ldots\}$ time units.

Sample Sequence. A drunk moves on a line with unit step forward with probab p and a unit step backward with probability $q = (p + q = 1)$. The format is

n	:	0	1	2	3	4	5	6	7	8	9	10 ...
Probab	:		p	q	q	p	p	p	q	q	p	q ...
$X(n)$	:	0	1	0	−1	0	1	2	−1	−2	−1	−2 ...

[The second row indicates his forward or backward steps and the third row indicates his position after his movements].

Now

$$E(Z_k) = (1)p + (-1)q = p - q, \; E(Z_k^2) = (1)^2 p + (-1)^2 q = p + q = 1$$

$$\text{Var}(Z_k) = E(Z_k^2) - E^2(Z_k) = 1 - (p - q)^2 = 4pq.$$

$$E(X_n) = E(\Sigma Z_k) = nE(Z_k) = n(p - q)$$

$$\text{Var}(X_n) = \text{Var}(\Sigma Z_k) = n\,\text{Var}(Z_k) = 4npq.$$

Note. When $p = q = 1/2$, $E(X_n) = 0$, Var $(X_n) = n$.

Comments. The simple RW, $X(n)$ has a Markov Property : Future is built on the present, not on the past. We show it

$$P\{Z_n = k\} = A_k, \; (k = 1, A_1 = p, k = -1, A_{-1} = q)$$

$$P\{X_{n+1} = i_{n+1} \mid X_0 = 0, X_1 = i_1, \ldots X_n = i_n\} = P\{Z_{n+1} + i_n = i_{n+1} \mid X_0 = 0, \; X_1 = i_1, X_n = i_n\}$$

$$= P\{Z_{n+1} = i_{n+1} - i_n\} = A_{i_{n+1} - i_n} = P\{X_{n+i} = i_{n+1} \mid X_n = i_n\}$$

because Z_{n+1} is independent of $X_0, X_1, \ldots, X_n$, these covering only $Z_1, \ldots, Z_n$. This shows that X_{n+1} (future) depends on X_n (present) and is independent of $X_{n-1}, X_1, \ldots, X_p, X_0$ (past).

13-12. *Worked-out Problems*

***Example* 1.** Given a Pois (λ), find the cumulative probabilties of (a) no occurrences (b) k or less occurrences. Write also expressions for $P(X \geq k)$ and hence show that

$$P(X \geq n) - P(X \geq n+1) = P(X = n)$$

Solution. Recall that $P(X = r) = f(r) = [e^{-\lambda} \lambda^r / r!]$, $r = 0, 1, 2, \ldots, \infty$.

(a) P [no occurences] $= P(X \leq 0) = P(X = 0) = e^{-\lambda}$.

(b) $$P(X \leq k) = P\left(\biguplus_{r=0}^{k} \{X = r\}\right) = \sum_{r=0}^{k} f(r) = e^{-\lambda} \sum_{r=0}^{k} \left(\frac{\lambda^r}{r!}\right) \quad \ldots(1)$$

(c) $$P(X \geq k) = P\left(\biguplus_{r=k}^{\infty} \{X = r\}\right) = \sum_{r=k}^{\infty} f(r) = e^{-\lambda} \sum_{r=k}^{\infty} \left(\frac{\lambda^r}{r!}\right) \quad \ldots(2)$$

or $$P\{X \geq k\} = 1 - P\{X < k\} = 1 - e^{-\lambda} \Sigma(\lambda^r / r!), r = 0, 1, \ldots k-1. \quad \ldots(3)$$

Replace k by n and $(n+1)$ in (2) to obtain

$$P(X \geq n) - P(X \geq n+1) = e^{-\lambda} \sum_{r=n}^{\infty} \left(\frac{\lambda^r}{r!}\right) - e^{-\lambda} \sum_{n+1}^{\infty} \frac{\lambda^r}{r!} = \frac{e^{-m} \lambda^n}{n!} = P(X = n). \quad \ldots(4)$$

Note. The result (4) is true for all discrete distributions.

Remarks. $P(X \geq 0) = 1 - P(X < 0) = 1, P(X \geq 1) = 1 - P(X < 1) = 1 - e^{-\lambda}$.

$$P(X \geq 2) = 1 - P(X < 2) = 1 - e^{-\lambda}(1 + \lambda), P(X \geq 3) = 1 - P(X \leq 2) = 1 - e^{-m}(1 + m + m^2/2!)$$

$$P(X \leq 1) = e^{-\lambda}(1 - \lambda), P(X \leq 2) = 1 - e^{-\lambda}(1 + \lambda + \lambda^2/2!), \text{etc.}$$

***Example* 2.** If $X \sim$ Pois (λ), show that $P(X \leq x) = \dfrac{1}{\Gamma(x+1)} \int_{\lambda}^{\infty} e^{-t} t^x \, dt, \quad x \geq 0.$ $\quad \ldots(1)$

Solution. Let $$I_k = \frac{1}{k!} \int_{\lambda}^{\infty} e^{-t} t^k \, dt. \qquad \text{[Incomplete gamma Integral]}$$

$$I_0 = \int_0^{\infty} e^{-t} \, dt = e^{-\lambda} = f(0)$$

Now we integrate by parts to obtain

$$I_k = \left[-\frac{t^k e^{-t}}{k!}\right]_{\lambda}^{\infty} + \int_{\lambda}^{\infty} \frac{e^{-t} t^{k-1}}{(k-1)!} dt = \frac{\lambda^k e^{-\lambda}}{k!} + I_{k-1}$$

$$\therefore \quad f(k) = I_k - I_{k-1}. \qquad [f(k) = e^{-\lambda} (\lambda)^k / k!]$$

$$\sum_{k=1}^{x} f(k) = \sum_{k=1}^{x} (I_k - I_{k-1}) = I_x - I_0 = I_x - f(0)$$

or $$F_x(x) = \sum_{k=0}^{x} f(k) = I_x \qquad \ldots\text{[Result (1)]}$$

Example 3. The number of events X in the interval $(0, t)$ is Pois (λt)-distributed but each event has only a probability p of being recorded. Find the distribution of the number of *recorded* events Y.

Solution. Recall : $P(X = r) = e^{-\lambda t}(\lambda t)^r / r!, \quad r = 0, 1, 2, \ldots$ We utilize multistage p-rule.

$$P\{Y = k\} = \sum_{r=k}^{\infty} P(Y = k \mid X = r)\, P(X = r) = \sum_{r=k}^{\infty} \binom{r}{k} q^{r-k} p^k \cdot \frac{e^{-\lambda t}(\lambda t)^r}{r!}$$

$$= (e^{-\lambda t}) \frac{(p\lambda t)^k}{k!} \sum_{r=k}^{\infty} \frac{(q.\lambda t)^{r-k}}{(r-k)!} \qquad \text{[Put } r - k = j\text{]}$$

$$= \frac{(e^{-\lambda t})(p\lambda t)^k}{k!} \left\{ \sum_{j=0}^{\infty} \frac{(q\,\lambda t)^j}{j} \right\} = \frac{e^{-\lambda t}.(p\lambda t)^k}{k!} e^{q\lambda t}.$$

$$= (e^{-p\lambda t})(p\lambda t)^k / k!$$

Thus $Y \sim$ Pois $(p\lambda t)$.

Problems with Solutions Provided at the End of the Text

1*. The r.v. X has Poisson distribution

(a) If $f(1) = f(2)$, find $f(4)$. (b) If $2f(0) + f(2) = 2f(1)$, find $\mathsf{E}(X)$.

(c) If $f(2) = 9f(4) + 90f(6)$, find $\mathsf{E}(X)$. (d) If $X \sim$ Pois (1), find $P(X \geq 2 \mid X \leq 4)$.

2*. If $X \sim$ Pois (λ), show that $P(X \geq 2) = \int_0^{\lambda} x e^{-x}\, dx$. ...(1)

3*. If Pois (λ) is so modified that the values $X = 0$, and $X = 1$ cannot occur, find the p.m.f. of the 2-pt truncated distribution.

4*. If X is Pois (λ) and k, r are positive integers such that $k \leq r$, prove that

$$\sum_{n=r}^{\infty} n^{(k)} P\{X = n\} = \lambda^k P\{X \geq r - k\}.$$

5*. If X is Pois (λ), show that $\lim_{r \to \infty} P\{X = r \mid X \geq r\} = 1, r = 0, 1, 2, \ldots$

6*. If X is Pois (λ), show that $\lim_{r \to \infty} P\{X = r \mid X \geq r\} = 1, r = 0, 1, 2, \ldots$

7*. If $X \sim$ Pois (λ) and $P(X \geq s) = \sum_{x=s}^{\infty} f(x)$, $S = \sum_{x=s}^{\infty} P(X \geq s)$, prove that

$$S = (1 - k)\, P(X \geq k) + \lambda P(X \geq k - 1).$$

8*. If X is Pois (λ) and discrete variate Y has the p.d.f.

$$P\{Y = r \mid X = x\} = \binom{x}{r} p^x q^{x-r}, 0 < p < 1, p + q = 1 \;\; 1,\; 0 \leq r \leq x.$$

Show that the unconditional distribution of Y is Pois (λp).

Exercise 13(a)

1. Let X be a Poisson r.v. Write $P(X=a)=f(a)$.
   (a) If $f(0)=f(1)=k$, show that $k=e^{-1}$.
   (b) If $f(1)=2f(2)$, find $f(k)$
   (c) If $f(0.01)=0.1, f(2)=0.2$ find $f(0)$
   (d) If $f(2)=\frac{2}{3}f(1)$, find $f(3)$..
   (e) If $f(0)=2f(1)$, find $P(X\geq 2)$.
   (f) $2f(2)=3f(1)+10f(0)$, find $P\{0\leq X<4\}$

   $\left[\textbf{Ans. } e^{-1}/k!, f(0)=(1/2)^{200/199}, (32/18)^{-4/3}, 1(3/2\sqrt{e})\right]$

2. If X is Pois (λ), $\lambda>0$ show that :
   (a) $P\{X=\text{even}\}=e^{-\lambda}\cosh\lambda=\frac{1}{2}(1+e^{-2\lambda})$.
   (b) $P\{X=\text{odd}\}=e^{-\lambda}\sinh\lambda=\frac{1}{2}(1-e^{-2\lambda})$.
   (c) $f(x)=f(x+1)$ for some integer $x\geq 0$, λ being an integer.

3. (a) If X is Pois (λ), find the distribution of $(-1)^X$.
   (b) Let $X\sim$ Pois (λ). If $f(x)=f(x+1)>0$, what value must λ have ?

   $\left[\textbf{Ans. } e^{-\lambda}\cosh\lambda, (X \text{ even});\ e^{-\lambda}\sinh\lambda\ (X \text{ odd}).\ \lambda=x+1\right]$

4. (a) Let $X\sim$ Pois (λ) and $F(n)=P(X\leq n)$. Show that

$$\sum_{r=0}^{n} rf(r) = \lambda F(n-1), \quad \sum_{r=0}^{n} r^2 f(r)=\lambda^2 F(n-2)+\lambda F(n-1).$$

   (b) Let $X\sim$ Pois (λ) and Y ~ U(0, 1) [$f(y)=1, 0\leq y\leq 1$]. Show :

$$P(X=0)=P(0<Y<e^{-\lambda}). \qquad P(X=n)=P\{F(n-1)<Y<F(n), n\geq 1.$$

5. Show that if $f(x)=(4/x)f(x-1)$, $x=1, 2, \ldots$ then $f(x)=e^{-4}4^x/x!\ 0\leq x<\infty$.

6. Let N be a fixed integer and $X>0$ be a constant. If X is Pois (λ), show that

$$\frac{1}{\lambda}P\{x+2\leq X\leq\alpha+1+N\}<\sum_{k=1}^{N}\frac{P\{X=\alpha+k\}}{\alpha+k}\leq\frac{(\alpha+2)}{(\alpha+1)}\frac{P\{\alpha+2\leq X\leq\alpha+1+N\}}{\lambda}.$$

7. If X is Pois (λ), F_X and f_x its c.d.f. and p.d.f., then for fixed x,
   (a) $f(x;\lambda)=(\lambda^x/x!)\,f(0,\lambda)\ [f(x,\lambda)\equiv f_x, F(x,\lambda)\equiv F_X]$
   (b) $df(x;\lambda)/d\lambda=f(x-1,\lambda)-f(x;\lambda)$
   (c) $xf(x;\lambda)=\lambda f(x-1;\lambda),\ \Sigma yf(y;\lambda)=\lambda F(x-1,\lambda),\ y=0,1,\ldots x$.

8. If X is a Poisson variate with parameter λ, and integer $r=0, 1, 2, \ldots$ show that

$$P(X>r) = \frac{1}{\Gamma(r+1)}\int_0^{\lambda} e^{-t}t^r\,dt.$$

9. Given that $g(x,0)=0$ and $\partial g(x,\omega)/\partial\omega=-\lambda g(x,\omega)+\lambda g(x-1,\omega)$ for $1\leq x<\infty$.

   If $g(0,\omega)=e^{-\lambda\omega}, \lambda>0$, using induction show that $g(x,\omega)=(\lambda\omega)^x e^{-\lambda\omega}/x!$.

10. Let X be Pois (λ). Show that $P(X=1\mid X\geq 1)=\lambda/(e^{\lambda}-1)$ and check whether this probability decreses with increasing λ.

11. Let X be Pois (λ). Find $p = P(X = k)/P(X \leq x)$ for positive integer k. Also show that p satisfies the differential equation : $\dfrac{dp}{d\lambda} = p\left[p + \dfrac{(k-\lambda)}{\lambda}\right]$.

12. Treat the Pois (λ) as a function of the parameter λ for fixed x. Show that f has maximum for $x = 1, 2, \ldots$, at $\lambda = x$ and f has two inflection points at $\lambda = x \pm \sqrt{x}$ for $x = 2, 3, \ldots$, and one inflection point at $\lambda = 2$ for $x = 1$.

13. If X pois (λ) show that $P(X > r) < \lambda^r / r\,!$, $r = 0, 1, 2, \ldots$ Deduce : $E(X) < e^{\lambda}$.

14. Let X and Y be two independent Pois (λ) variates. Find the p.m.f of max (X, Y) and min (X, Y).

$$\left[\textbf{Ans. } f(z)\,[2F(z-1) + f(z)]\,[2 - f(z)\,[2 - F(z) - F(z-1)],\ f(z) = e^{-\lambda}\,\lambda^z / z\,!\right]$$

15. Let $X_1, X_2, \ldots$ be a sequence of independent bin $(1, p)$ variates with $0 < p < 1$. Also, let $S_N = X_1 + X_2 + \ldots + X_N$, where $N \sim$ Pois (λ). Show that S_N and $N - S_N$ are independent.

16. Let X and Y be independent r.v.s. with $P(X = k) = p_k$ $P(Y = k) = q_k$ $0 \leq k < \infty$ and $\Sigma p_k = 1 = \Sigma q_k$. Let $P(X = k \mid X + Y = t) = \binom{t}{k} (\lambda_t)^k\,(1 - \lambda_t)^{t-k}$, $0 \leq k \leq t$.

Then $\lambda_t = \alpha\ \forall t$, and $p_k = \dfrac{e^{-\beta\theta}(\beta\theta)^k}{k\,!}$, $q_k = \dfrac{e^{-\theta}\,\theta^k}{k\,!}$ where $\beta = \dfrac{\alpha}{1-\alpha}$, and $\theta > 0$ is arbitrary.

17. Let $X \sim$ Pois (λ) but $\lambda \sim$ expo (1). Show that $P(X = n) = (1/2)^{n+1}$.

18. Construct a Poisson counting process based on the definition of Poisson p.m.f. and establish its equivalence with §13-11.

13-20. *Poisson Approximation to the Binomial pmf*

If $p \to 0$ and $n \to \infty$ in such a manner that the mean np $= \lambda$ (say) > 0, remains fixed, then the bin (n, p) density approaches the Pois (λ) density.

Proof.

$$f(x) = \binom{n}{x} q^{n-x} p^x\ \frac{n(n-1)(n-2)\ldots(n-x+1)}{x\,!}\ p^x (1-p)^{n-x}$$

$$= \frac{1}{x\,!}\left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right)\ldots\left(1 - \frac{x-1}{n}\right)(np)^x\,(1-p)^{n-x} \qquad \ldots(1)$$

Recall : $\lim\,(1 + a/n)^{bn+c} = e^{ab}$, $(n \to \infty)$ (Euler's Limit) ...(A)

$$\lim_{n\to\infty}(1-p)^{n-x} = \lim_{n\to\infty}\left(1 - \frac{\lambda}{n}\right)^{n-x} = e^{-\lambda} \quad \left(\because\ p = \frac{m}{n}\right)$$

Also $\displaystyle\lim_{n\to\infty}\left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right)\ldots\left(1 - \frac{x-1}{n}\right) = (1-0)\,(1-0)\ldots(1-0) = 1$

Making substitutins into (1) we get

$$\lim f(x) = e^{-\lambda}\lambda^x / x\,!,\ x = 0, 1, 2, \ldots \infty.$$

Note. We can also obtain a Poisson approximation to the binomial distribution for the case when n is large and nq is small. Letting $y = n - x$, we have

$$\lim_{n\to\infty}\binom{n}{x} p^x q^{n-x} = \lim_{n\to\infty}\binom{n}{n-x}(1-q)^x\,q^{n-x} = \lim_{n\to\infty}\binom{n}{y} q^y\,(1-q)^{n-y} = e^{-\lambda}\lambda^y / y\,!.$$

where $\lambda = nq$ is the paramter of the Poisson Distribution.

***Example* 1.** Let $X_1, X_2, ..., X_n$ be i.i.d bin $(1, p)$ variates Let $S_n = X_1 + ... + X_n$ and $M_n(t)$ be the m.g.f. of S_n. Find $\lim M_n(t)$ using $np = \lambda$ (const.)

Solution. Here $M(t : X_1) = (q + pe^t)^1$, so $M(t : S_n) = [M(t : X_1)]^n = (q + pe^t)^n$.

Thus, $$M_n(t) = [1 = p(e^t - 1)]^n = [1 + (e^t - 1)\lambda / n]^n$$

$$\lim_{n\to\infty} M_n(t) = \lim_{n\to\infty}\left[1 + \frac{(e^t - 1)\lambda}{n}\right] = e^{\lambda(e^t - 1)}. \quad \text{[By Euler's limit]}$$

It follows that $S_n \to$ Pois (λ) as $n \to \infty$, with $np = \lambda$ (fixed).

***Example* 2.** Let $X_1, X_2, ..., X_n$ be independent Indicator (Bernoulli) variables : $P(X_j = 1) = p_j, P(X_j = 0) = q_j$ and $S_n = X_1 + X_2 + ... + X_n$. Show that if $\max p_j \to 0$ and $(p_1 + p_2 + ... + p_n) \to \lambda$, $(0 < \lambda < \infty)$ as $n \to \infty$ then $S_n \to$ Pois (λ).

Solution. $G(t : X_j) = \mathrm{E}(t^{X_j}) = q_j t^0 + tp_j = 1 + (t-1)p_j = 1 + t' p_j \quad (t' = 1 - t)$

$$\therefore \quad G_n(t) = G(t : S_n) = G(t : X_1 + ... + X_n) = G(t : X_1).G(t : X_2) ... G(t : X_m) = (1 + t'p_1) ... (1 + t'p_n) \quad ...(1)$$

$$\therefore \quad \ln G_n(t) = \sum_{j=1}^{n} \ln(1 + t'p_j) = \sum_{j=1}^{n}\left[(p_j t') - \frac{(p_j t')^2}{2} + \frac{(p_j t')^3}{3} - ...\right]$$

As $n \to \infty$, $\Sigma p_j \to \lambda$, $\Sigma p_j^k \to 0$ (as $\max p_j \to 0$, $k = 2, 3, ...$) and thus

$$\ln G_n(t) \to \lambda t' \quad \Rightarrow \quad G_n(t) \to e^{\lambda(t-1)}.$$

This is p.g.f. of Pois (λ). Hence by Continuity Theorem, $S_n \to$ Pois (λ).

Remark. If $p_j = p, \forall_j$ then (1) provides :

$$G_n(t) = (1 + t'p)^n = [1 + (\lambda t'/n)]^n.$$

Thus $G_n(t) \to e^{\lambda(t-1)}$: Poisson approximation to the Binomial distribution vide Example 1.

Occurrence of Poisson Distribution

The Poisson distribution finds applications in a wide variety of situations in which some kind of *event* occurs repeatedly, but *haphazardly.*

Typically situtations are those in which the events of the following nature occur :

(a) Number of telephone calls arriving at an exchange in some unit of time.

(b) The number of defective items in a given lot.

(c) The emission of radio-active (alpha) particles.

(d) Completion of a repair-job by repairmen.

(e) Number of defects in a tape or in a sheet of manufactured material.

(f) The number of printing (typing errors.

(g) Bacterial colony on a Petri plate.

Treated as an example of a rare event (Art 13-20), such situations as under are encountered :

1. The number of rail road (or air-crashes) accidents in some unit of time.
2. The number of insurance claims in some unit of time.

3. The number of deaths caused by the kick of a horse.
4. The number of deaths caused by suicide or by diseases such as snake-bite, heart-attack, cancer, etc. in some unit of time, etc.

13-21. *Worked-out Numericals*

***Example* 1.** A certain hospital usually admits 50 patients per day. On the average 3 patients in 100 require special facilities found in special rooms. On the morning of a certain day, it is found that there are three such rooms available. Assuming that 50 patients will be admitted, find the probability that more than three patients will require such special rooms.

Solution. Let X be the r.v. denoting the number of patients that will require the special rooms. Then X is bin (n, p), where $n = 50$, $p = 0.03$. We need find $P(X > 3)$. Now

$$P(X > 3) = 1 - P(X \leq 3) = 1 - \sum_{x=0}^{3} \binom{50}{x} (0.03)^x (0.97)^{50-x}.$$

We can approximate this sum by Poisson distribution. Here $m = np = 50 \times 0.03 = 1.5$; hence

$$P(X > 3) = 1 - P(X \leq 3) = 1 - \sum_{x=0}^{3} \frac{e^{-1.5}(1.5)^x}{x!} = 1 - 0.934 = 0.066.$$

***Example* 2.** Assume that the number of cars passing an intersection obeys a Poisson distribution. If the probability of no cars in 1 minute is 0.20, what is the probability of more than one car in 2 minutes ?

Solution. $P(X = 0) = e^{-\lambda} = 0.20$; hence $\lambda = -\ln 0.2 = \ln_e 5 = 1.60944$.

Now taking 1 minute as a unit of time, we have $\lambda = 1.60944$. Hence the average number of arrivals per 2-minute interval is $2m$. Now if $X \sim$ Pois (2λ), then

$$\begin{aligned} P(X > 1) &= 1 - P(X \leq 1) = 1 - [P(X = 0) + P(X = 1)] 1 - [e^{-2\lambda} + e^{-2\lambda}(2\lambda)] \\ &= 1 - e^{-2\lambda}(1 + 2\lambda) = 1 - (0.2)^2 [1 + 3.2188] = 1 - 0.168752 = 0.8312448. \end{aligned}$$

Problems with Solutions Provided at the End of the Text

1*. If the probability that a individual suffers a bad reaction from injection of a given serum is 0.001, find the probability that out of 2000 individuals, (i) exactly 3, (ii) more than 2, individuals will suffer from a bad reaction.

2*. It is known tha the probability that an item, produced by a certain machine will be defective is 0.01. By applying Poisson's approximations, show that the probability that the random sample of 100 items selected at random from the total output will contain no more than one defective item is $2/e$.

3*. A device contains 2000 equally reliable elements with prob. of failure for each of them equal to 0.0005. What is the probability that the device will fail to operate if failure occurs when at least one element fails to operate ?

4*. The King's minister boxes his coins n to a box. Each box contains k false coins. The King doubts the minister and randomly draws a coin from each of n boxes and gets them tested. Find the chance that

(i) King's sample of n coins contains exactly r false coins.

(ii) Minister's peculations go undetected.

5*. The proofs of a 500-page book contain 500 misprints. Find the prob. that there are at least four misprints per page.

6*. An insurance company finds that 0.005 percent of the population die of road-crossing each year. What is the prob. that the company must pay off on more than 3 of 10,000 insured risks, against such accidents in a given year ?

7*. Obtain a formula for the number of Poisson trials required for a specified certainty of at least one success.

If intact specimens occur on the average of 2 per 10,000 cubic feet of earth removed, how many cubic feet of earth is to be removed in order to be 95% sure of obtaining an intact specimen ?

8*. The number of criminals being hanged in a week is a Poisson variate with mean λ. However, λ itself is a variate which takes on one of the values 1, 2, 3, 4 with respective probablities 1/3, 1/3, 1/4, 1/12. What is the probability that no criminal is hanged.

9*. The number of cars passing an intersection during any time interval of length t minutes between 6 P.M. and 7 P.M. is Pois (t) distribution. Let T be time elapsed after 6 P.M. before the first car crosses the intersection. Find $P(T < 1)$.

13-30. *Two-Parameter Poisson Variate*

In many applications of Poisson variate, it is convenient to replace λ by ct and obtain

$$f(x\,;\,c,\,t) = e^{-ct}(ct)^x\,/\,x!,\;\; c>0,\; t>0,\; x=0,1,2,\ldots;\; f(x;c,t)=0,\text{ otherwise} \qquad \ldots(A)$$

We call $f(x\,;\,c,\,t)$ the Poisson distribution with parameters c and t. Here t is usually an interval of time (or space) over which the number of occurences of some event is observed, c is the average number of such occurences per unit time (space). $f(x\,;\,c,\,t)$ gives the prob.of exactly x occuences in an arbitrary interval of time of length t.

Example : On the average, a submarine on patrol sights 6 enemy ships per hour. Assuming the number of ships sighted in a given length of time is Poisson, find the prob. of sighting

(a) 6 ships in the next half hour,

(b) 4 ships in the next two hours.

(c) at least one-ship in the next 15 min.,

(d) at least two ships in the next 20 min

Solution. (a)Here $c = 6$, $t = 1/2$, ct $= \lambda = 3$

$$P(X=6) = \frac{e^{-3}(3)^6}{6!} = \frac{0.0498\times 9\times 9\times 9}{720} = 0.05042.$$

(b) Here $c = 6$, $t = 2$, $\lambda = ct = 12$

$$\therefore \quad P(X=4) = \frac{e^{-12}(12)^4}{4!} = (0.0025)^2\,144.6 = 0.0054.$$

(c) Here $c = 6$, $t = 1/4$, $\lambda = ct = 3/2 = 1.5$

$$P(X\ge 1) = 1-P(X=0) = 1-e^{-1.5} = 1-0.2231 = 0.7769.$$

(d) Here $c = 6$, $t = 1/3$, $\lambda = ct = 2$

$$P(X \geq 2) = 1 - [P(X = 0) + P(X = 1)] = 1 - 3e^{-2} = 1 - 3 \times 0.1353 = 1 - 0.4059 = 0.5941.$$

Exercise 13(b)

1. Use Poisson distribution in order to approximate the binomial distribution :
   (a) $n = 100, p = 0.02, x = 5$ (b) $n = 20, p = 0.05, x = 5$
2. Six coins are tossed 6400 times, Using the Poisson distribution, show that the approximate value of getting 6 heads x times is $e^{-100} . (100)^x/x\,!$.
3. A manufacturer of cotter pins knows that 5% of his product is defective. If he sells cotter pins in boxes of 100 and gurantees that not more than 10 pins will be defective, show that the approximate probability that a box will fail to meet the guranteed quality is 0.013695.
4. A distributor of bean seeds determines from extensive tests that 5% of large batch of seeds will not germinate. He sells the seeds in packets of 200 and gurantees 90% germination. Determine the probability that a particular packet will voilate the guarantee.
5. Patients arrive randomly and independently at a doctor's Surgery from 8 A.M. at an average rate of one in five minutes. the waiting room holds 12 persons. Show that the probability that the room will be full wen the doctor arrives at 9 A.M. is 0.5384. (Estimate the probability to an accuracy of 5 percent).
6. A car hire firm has two cars, which it hires out day by day. The number of demands for a car on each day is Poisson distributed with mean 1.5. Calculate the proportion of days on which (a) neither car is used. (b) some demand is refused.($e^{-1.5} = 0.2231$) [**Ans.** 0.2331 ; 0.19126]
7. A man is standing some distance from blasting operations in a quarry. Ten small pieces of debris fall randomly within space of 12 sq. yd. If the plan area of a man is 2 sq. ft. show that the probability that if the were standing within 12 sq. yd, he would not have been hit by any of the pieces of debris is $e^{-5/27}$.
8. A telephone switchboard handles 600 calls, on an average, during a rush hour. The board can make a maximum of 20 connections per minute. Show that the probability that the board will be over-taxed during any given minute is $\Sigma e^{-10}\,(10)^x\,/\,x!,\ 21 \leq x < \infty$,
9. The average number of telephone calls reaching the switch board of Daffodil-Company is 30 calls per hour. Find the probability that (a) No call arrives in a 3-minute period (b) More than five calls arrive in a 5 minute period.
10. The number of telephone calls that an operator receives from 9.00 to 9.05 hours in a day follows Poisson distribution with mean 3. Show that the probability that
    (i) The operator will receive no calls in that time interval tommorrow is e^{-3}.
    (ii) In the next three days the operator will receive a total of 1 call in that time interval is $9e^{-9}$.
11. A radioactive source is observed during 4 time intervals of a 6 seconds each. If the number of particles emitted during each time interval follows a Poisson probability law with a rate of 0.5 particles per second, show that the probability that (a) in each of the four time intervals 3 or more particles will be emitted, is 0.109.
    (b) in atleast one of the four time intervals 3 or more particles will be emitted, is 0.968.
12. A large number of observations on a given solution which contained bacteria were made taking samples 1 ml each. nothing down the number of bacteria present in each sample. Assuming Poisson distribution and given that 10% samples contained no bacteria, show that the average number of bacteria per ml is $\ell n_e\ 10$.

13. Red blood cell deficiency may be determined by examining a specimen of the blood under a micro-scope. Suppose that a certain small fixed volume contains, on the average, 20 red cells for normal person. Using Poisson distribution, show that the probability that a specimen from a normal person will contain less than 15 red cell is $\Sigma e^{-20}(20)^x / x!,\ 0 \le x \le 14$.

14. (a) The chance of a traffic accident in a day, attributed to taxi drivers is 0.001. Out of a total of 1000 days, on how many days we shall have

(i) no accident. (ii) more than 10 accidents. [**Ans.** (i) 368 days]

(b) You sell an average of 2 T-V tubes per day. How many tubes would you stock so that, with probability that at least 0.95, you will meet demand for a 30-day month ? (Sales obey Poisson law). Generalize the sales for time-period T, with λ as the mean rate of occurence per unit time.

15. In a double sampling scheme, a sample of 50 items is selected from a batch, and the batch is accepted if the sample has no defectives. If the sample hsa 3 or 4 defectives, the batch is rejected. If the sample has one or two defectives, a second sample of 100 items is taken and the total number of defectives in the two samples is noted. If this total of defectives is five or more, the batch is rejected, otherwise the batch is accepted. Show that the probability of accepting a batch with 1% defective items is 0.974. What is the average of inspection per batch ?

13-40. *Recurrence formula for Poisson Density*

When computing the Poisson density numerically, one can often use a process of iteration based on the equations

$$f(x+1) = [\lambda/(x+1)].f(x); \qquad [f(0)=e^{-\lambda}]. \qquad \text{...(A)}$$

To prove it, we notice that

$$\frac{f(x+1)}{f(x)} = \frac{e^{-\lambda}\lambda^{x+1}/(x+1)!}{e^{-\lambda}\lambda^{x}/x!} = \frac{\lambda}{x+1}.$$

This yields (A). $f(0) = e^{-\lambda}$, is first evaluated and then (A) is applied recursively to compute the remaining $f(x)$, one after the other.

***Example* 1.** A Poisson distribution has a mean of 0.63. Find respective probabilities of 0, 1, 2, and 3 occurence of the event in question.

Solution. $f(0) = e^{-\lambda} = e^{-0.63} = 0.5326$ [by def.]

Putting $x = 0, 1, 2, 3$ sucessively in the Recurrence formula we get

$$f(1) = \lambda.f(0) = (0.630)(0.5326) = 0.3355;\ f(2) = \tfrac{1}{2}\lambda f(1) = (0.315)(0.3355) = 0.1055$$

$$f(3) = \tfrac{1}{2}\lambda f(2) = (0.210)(0.1055) = 0.0222.$$

***Example* 2.** In a certain factory turning razor blades, there is a small chance 1/500 for any blade to be defective. The blades are supplied in a packet of 10. Use Poisson distribution to calculate the approximate number of packets containing no defective, one defective, two defective, three defective and four defective blades, in a consignment of 20,000 packets.

Solution. Here $n = 10$, $p = 1/500$ so that $\lambda = np = 0.02$. If $f(x)$ is the frequency of x defective packets in $N = 20{,}000$ packets then

$f(x) = Ne^{-\lambda}\lambda^x / x! = 20{,}000 \times 0.9802 \times (0.02)^x / x! \left[f(x+1) = \frac{\lambda}{x+1} f(x) \right]$

$\therefore$ $f(0) = 19604$ (Number of packets with no defective blades).

So $f(1) = (0.02) \times 19604 = 392.08 = 392$; $f(2) = 0.01 \times 392.08 = 3.9208 = 3$

$f(3) = (0.02 \times 3.9208)/3 = 0.026137$. Obviously $f(4)$ is nearly zero.

Problems with Solutions Provided at the End of the Text

1*. Find mean and variance of Pois (λ) by differentiating the exponential series.

2*. Critize the following statement :

"The mean of a Poisson distribution is 5, while S.D. is 4".

3*. If X is Poisson distributed with parameter 100, use Chebyshev inequality to determine a lower bound for $P\{75 < X < 125\}$.

13-41. *Mode of the Poisson Distribution*

We have to find the value of the random variable $X \sim$ Pois (λ) for which the prob. is maximum. Thus if x is the modal value its definition provides : (i) $f(x) \geq f(x+1)$ (ii) $f(x) \geq f(x-1)$

The Poisson Recurrence formula is : $f(x+1) = [\lambda/(x+1)]\, f(x)$. Using this we see that

$$f(x+1) \leq f(x) \Rightarrow x+1 \leq \lambda, \quad i.e. \quad x \leq \lambda - 1. \quad f(x) \geq f(x-1) \quad \Rightarrow x \geq \lambda.$$

Combining the last two inequations we obtain

$$\lambda - 1 \leq x \leq \lambda \qquad \text{...(1)}$$

***Case* 1.** Let λ be a positive integer. Since the integer x lies between two consective integers $\lambda - 1$ and λ, equation (1) is possible iff $x = m$ and $x = m - 1$. And thus there are two modes for the distribution in this case.

***Case* 2.** Let λ be a fraction (rational number). Then the integer x lies between two fractions differing by 1. Hence x must be the integeral part of λ written $[\lambda]$.

***Example* 1.** A Poisson distribution has a double mode at $x = 1$ and $x = 2$. What is the probability that X will have one or the other of these two values.

Solution. Poisson's mode is given by $\lambda - 1 \leq x \leq \lambda$. Hence when $x = 1$, $\lambda - 1 = x$ gives $\lambda = 2$, and $x = 2$, $\lambda = x$ gives $\lambda = 2$. Now $P(X = k) = e^{-\lambda}\lambda^k/k\,!$.

$$\therefore \qquad P\{X = 1\} = \frac{e^{-2}\,2!}{1!} = \frac{2}{e^2}; \quad P\{X = 2\} = \frac{e^{-2}(2)^2}{2!} = \frac{2}{e^2}.$$

Hence $\quad P\{(X = 1) \uplus (X = 2)\} = P\{X = 1) + P(X = 2) = 4/e^2$.

Exercise. At a certain telephone exchange the average number of calls passed per hour in the morning is 120 and the rate can be regarded as constant. Calculate the probability of exactly 3 calls in a period of 5 minutes, and of more than 3 calls in the same period. What is the most likey number of calls ?

***Example* 2.** Monotonicity. For Pois (λ) let $f(k) = e^{-\lambda} m^k / k!$, $0 \le k < \infty$. Show that $f(k-1) >, =, < f(k) \Leftrightarrow k >, = < m$.

Solution.
$$\frac{f(k-1)}{f(k)} = \frac{e^{-\lambda} \lambda^{k-1} / (k-1)!}{e^{-\lambda} \lambda^k / k!} = \frac{k}{m}.$$

$\therefore \quad f(k-1)/f(k) >, =, <1 \Leftrightarrow (k/m) <, = < 1, \Rightarrow f(k-1) >, =, < f(k) \Leftrightarrow k >, =, < \lambda.$

Note. $f(k-1) = f(k)$ holds if λ is an integer.

13.42. *Mean (absolute) Deviation (M.a.D)*

If X is Poisson (λ), then by definition

$$M = \mathsf{E}\,|X - \lambda| = \sum_{x=0}^{\infty} f(x)\,|x-\lambda| = \sum_{x>\lambda} f(x)(x-\lambda) + \sum_{x<\lambda} f(x)(\lambda - x).$$

$$0 = \mathsf{E}(X-\lambda) = \sum_{x=0}^{\infty} f(x)(x-\lambda) = \sum_{x>\lambda}(x-\lambda) + \sum_{x<\lambda} f(x)(x-\lambda).$$

Adding these relations, using $\theta = [\lambda + 1]$: greatest integer function

$$\frac{1}{2}M = \sum_{x>\lambda} f(x)(x-\lambda) = e^{-\lambda} \sum_{x=\theta}^{\infty}\left[\frac{\lambda^x}{(x-1)!} - \frac{\lambda^{x+1}}{x!}\right]. \qquad \left[f(x) = \frac{e^{-\lambda}\lambda^x}{x!}\right]$$

$$= e^{-\lambda} \sum_{x=\theta}^{\infty} (T_x - T_{x+1}) = e^{-\lambda} T_\theta. \quad \text{(Telescopic Series)} \qquad \left[T_x = \frac{\lambda^x}{(x-1)!}\right]$$

$$\therefore \qquad M = 2e^{-\lambda} . \lambda^\theta / (\theta - 1)!, \text{ where } \theta = [\lambda + 1]$$

Note. $\lim\limits_{n\to\infty} \dfrac{\lambda^x}{x!} = \lim\limits_{n\to\infty} T_x = 0$, since this series is convergent.

***Example* 1.** Show that if X is Pois (1) its mean deviation about mean is $(2\sigma_X / e)$.

Solution. The result is immediate from §13-42, putting $\lambda = 1$. However, we may proceed independently :

$$M = \mathsf{E}(|X-\lambda|) = \sum_{x>\lambda}(x-\lambda) f(x) + \sum_{x<\lambda}(\lambda - x) f(x),$$

$$0 = \mathsf{E}(X-\lambda) = \sum_{x>\lambda}(x-\lambda) f(x) + \sum_{x<\lambda}(x-\lambda) f(x).$$

Adding these equations and using $\lambda = 1$, we get

$$\frac{1}{2}M = \sum_{x>1}(x-1) f(x) = \sum_{x=2}^{\infty} x f(x) - \sum_{x=2}^{\infty} f(x) = \left[\sum_{x=0}^{\infty} x f(x) - f(1)\right] - \left[\sum_{x=0}^{\infty} f(x) - f(0) - f(1)\right]$$

$$= [\lambda - f(1)] - [1 - f(0) - f(1)] = f(0) = e^{-1} \qquad [f(x) = e^{-\lambda}\lambda^x / x!;\ \lambda = 1]$$

$\therefore \quad M = (2/e) = (2\sigma_X / e)$ as $\sigma_X^2 = \lambda = 1$

Alter. $M = \text{E}\,|X - 1| = \sum_{x=0}^{\infty} f(x)\,|x-1| = f(0) + \sum_{2}^{\infty} f(x)\,(x-1) = f(0) + \sum_{x=2}^{\infty} x\,f(x) - \sum_{x=2}^{\infty} f(x).$

$$= f(0) + \left[\sum_{x=0}^{\infty} x\,f(x) - f(1)\right] - \left[\sum_{x=0}^{\infty} f(x) - f(0) - f(1)\right] = 2f(0) + \lambda - 1 = 2/e. \quad [\lambda = 1]$$

Example 2. If X is Pois (λ), show that $\text{E}(|X-1|) = \lambda + 2e^{-\lambda} - 1.$

Solution. Here $\qquad f(x) = e^{-\lambda}\lambda^x / x!, \quad x = 0, 1, 2, \ldots$

$$\begin{aligned} \text{E}(|X-1|) &= \sum_{x=0}^{\infty} f(x)\,.\,|x-1| = f(0) + \sum_{x=2}^{\infty} f(x)\,.\,(x-1) = f(0) + \sum_{x=2}^{\infty} x\,f(x) - \sum_{x=2}^{\infty} f(x) \\ &= f(0) + \left[\sum_{x=0}^{\infty} x\,f(x) - f(1)\right] - \left[\sum_{x=0}^{\infty} f(x) - f(0) - f(1)\right] = 2f(0) + \sum_{x=0}^{\infty} x\,f(x) - \sum_{x=0}^{\infty} f(x) \\ &= 2e^{-\lambda} + \lambda - 1. \qquad [f(0) = e^{-\lambda}] \end{aligned}$$

13-50. *Moments of Poisson Distribution*

Direct derivation of moments. The following chain is very convenient to obtain Poisson's moments.

Simple factorial moments $\to$ Simple moments $\to$ Central moments.

Simple factorial moments :

$$\mu'_{(r)} = \text{E}(X^{(r)}) = \sum_{x=0}^{\infty} \frac{e^{-\lambda}\lambda^x}{x!}\,x^{(r)} = e^{-\lambda}\lambda^r \sum_{x=r}^{\infty} \frac{\lambda^{x-r}}{(x-r)!} = (e^{-\lambda}\,\lambda^r)\,e^{\lambda} = \lambda^r.$$

Thus, $\mu'_{(r)} = \lambda^r$. In particular, $\mu'_{(1)} = \mu = \lambda,\ \mu'_{(2)} = \lambda^2$, etc.

Simple moments through simple factorial moments :

$$\begin{aligned} \mu'_2 &= \text{E}(X^2) = \text{E}\{X^{(2)} + X\} = \text{E}(X^{(2)}) + \text{E}(X) = m^2 + m. \\ \mu'_3 &= \text{E}(X^3) = \text{E}\{X^{(3)} + 3X^{(2)} + X\} = \lambda^3 + 3\lambda^2 + \lambda \\ \mu'_4 &= \text{E}(X^4) = \text{E}\{X^{(4)} + 6X^{(3)} + 7X^{(2)} + X\} = \lambda^4 + 6\lambda^3 + 7\lambda^2 + \lambda. \end{aligned}$$

Central moments through simple moments :

$$\begin{aligned} \mu_2 &= \mu'_2 - (\mu'_1)^2 = (\lambda^2 + \lambda) - \lambda^2 = \lambda \\ \mu_3 &= \mu'_3 - 3\mu'_2\,\mu_1 + 2(\mu'_1)^3 = (\lambda^3 + 3\lambda^2 + \lambda) - 3(\lambda^2 + \lambda)\lambda + 2\lambda^3 = \lambda \\ \mu_4 &= \mu'_4 - 4\mu'_3\,\mu'_1 + 6\mu'_2(\mu'_1)^2 - 3\mu'^4_1 \\ &= (\lambda^4 + 6\lambda^3 + 7\lambda^2 + \lambda^2) - 4\lambda(\lambda^3 + 3\lambda^2 + \lambda) + 6\lambda^2(\lambda^2 + \lambda) - 3\lambda^4 = 3\lambda^2 + \lambda \end{aligned}$$

Ramanvosky Moments Recurrence Formula

$$\mu_{r+1} = \lambda r\,\mu_{r-1} + \lambda\,(d\mu_r / d\lambda).$$

Proof. By definition, $\mu_r = \text{E}(X-\lambda)^r = \sum_{x=0}^{\infty} f(x)\,(x-\lambda)^r = \sum_{x=0}^{\infty} \frac{e^{-\lambda}\lambda^x}{x!}\,(x-\lambda)^r.$

We assume that the above series is uniformly convergent. Now

$$\frac{d}{dm}\mu_r = \sum_{x=0}^{\infty}\frac{1}{x!}\frac{d}{dm}[(e^{-\lambda}\lambda^x)(x-\lambda)^r]$$

$$= \sum_{x=0}^{\infty}\frac{1}{x!}[(x\lambda^{x-1}e^{-\lambda}-e^{-\lambda}\lambda^x)(x-\lambda)^r - r(e^{-\lambda}\lambda^x)(x-\lambda)^{r-1}]$$

$$= \sum_{x=0}^{\infty}\frac{e^{-\lambda}\lambda^x}{x!}\left\{\left(\frac{x}{\lambda}-1\right)(x-\lambda)^r - r(x-\lambda)^{r-1}\right\} = \sum_{x=0}^{\infty} f(x)\left[\frac{(x-\lambda)^{r+1}}{\lambda} - r(x-\lambda)^{r-1}\right]$$

$$= \mathrm{E}[\lambda^{-1}(X-\lambda)^{r+1} - r(X-\lambda)^{r-1}] = \lambda^{-1}\mathrm{E}(X-\lambda)^{r+1} - r\mathrm{E}(X-\lambda)^{r-1} = \lambda^{-1}\mu_{r+1} - r\mu_{r-1}.$$

Thus, by transfer we obtain

$$\mu_{r+1} = \lambda[r\mu_{r-1} + d\mu_r/d\lambda]$$

Note. Putting $r = 1, 2, 3$ in succession, using $\mu_0 = 1$, $\mu_1 = 0$, $D = d/d\lambda$ obtain

$$\mu_2 = \lambda\mu_0 + \lambda D\mu_1 = \lambda, \quad \mu_3 = 2\lambda\mu_1 + \lambda D\mu_2 = \lambda D(\lambda) = \lambda, \quad \mu_4 = 3\lambda\mu_2 + \lambda D\mu_3 = 3\lambda^2 + \lambda.$$

M.G.F., P.G.F., Moments and Cumulants

$$M_X(t) = e^{\lambda(e^t-1)}, G_X(t) = e^{\lambda(t-1)} \qquad [\S\ 8\text{-}16(3)\quad \S\ 8\text{-}51(2)]$$

We can find mean or Poisson distribution at once. Thus

$$\mu_1' = M'_{(0)} = [\lambda e^t . e^{\lambda(e^t-1)}]_{t=0} = \lambda$$

To find central *m.f.g.* we have

$$M(t : X-\lambda) = e^{-\lambda t}M(t:X) = e^{-\lambda t}.e^{\lambda(e^t-1)} = e^{\lambda}(e^t - 1 - t).$$

We now attend to cumulants :

$$K(t) = \ln M(t) = \ln[e^{\lambda(e^t-1)}] = \lambda(e^t-1) = \lambda[t + (t^2/2!) + (t^3/3!) + (t^4/4!) + (t^5/5!) + ...]$$

Since k_r = coefficient of (t^r/r) in $K(t)$, we immediately obtain

$$k_1 = \lambda, \qquad k_2 = \lambda, ..., k_r = \lambda.$$

Thus all the cumultants of Pois (λ) equal the parameter λ (mean) of the distribution. Since $k_r = \lambda\ \forall\ r$, we need no recurrence formula for k_r.

Obviously : $\mu = k_1 = \lambda, \qquad \mu_2 = k_2 = \lambda, \qquad \mu_3 = k_3 = \lambda, \qquad \mu_4 = k_4 + 3k^2 = \lambda + 3\lambda^2.$

$$\beta_1 = \frac{(\mu_3)^2}{(\mu_2)^3} = \frac{\lambda^2}{\lambda^3} = \frac{1}{\lambda}; \qquad \beta_2 = \frac{\mu_4}{(\mu_2)^2} = \frac{\lambda + 3\lambda^2}{\lambda^2} = 3 + \frac{1}{\lambda}.$$

$$\gamma_1 = \sqrt{\beta_1} = \lambda^{-1/2}, \qquad \gamma_2 = \beta_2 - 3 = \lambda^{-1}; \qquad \mu\sigma\gamma_1\gamma_2 = 1.$$

Example : If $M_X(t) = \exp\left[\sum_{r=1}^{\infty}\frac{t^r}{r!}\right]$, find $P(X > 2)$.

Solution.
$$M_X(t) = \exp\left[\sum_{r=0}^{\infty}\frac{t^r}{r!} - 1\right] = \exp(e^t - 1).$$

This is the m.g.f. of Pois (1) ; hence the p.m.f. of X is

$$f(x) = P(X = x) = e^{-1}/x\,!, \quad x = 0, 1, 2, \ldots$$

$$\therefore \quad P(X > 2) = 1 - P(X \le 2) = 1 - [f(0) + f(1) + f(2)] = 1 - e^{-1}[1 + 1 + (1/2)] = 1 - (5/2e).$$

First Four Central Factorial Moments

For Pois (λ), Factorial moments are $\mu'_{(r)} = \lambda^r$ [§ 8-31(2)]. Also using relations between $\mu_{(r)}$ and μ_r we obtain

$$\mu_{(2)} = \mu_2 = \lambda; \quad \mu_{(3)} = \mu_3 - 2\mu_2 = \lambda - 3\lambda = -2\lambda.$$

$$\mu_{(4)} = \mu_4 - 6\mu_3 + 11\mu_2 = (2\lambda^3 + \lambda) - 6\lambda + 11\lambda = 3\lambda^2 + 6\lambda = 3\lambda(\lambda + 2).$$

13-60. *Additive Property. Subtractive Property*

1. If X_1 and X_2 are two independently distributed Poisson variates with paramters λ_1 and λ_2 respectively, then $X_1 + X_2$ is also a Poisson variate with parameter $\lambda_1 + \lambda_2$. [Proof. § 8-54(2)]

Comments. The result is not true if X_1 and X_2 are dependent.For instance, take $X_2 = X_1$, then $X = X_1 + X_2 = 2X_1$ takes only even values, hence it cannot have the Poisson distribution.

2. If $X_i \sim$ Pois (λ_i) ; X_1 and X_2 are independent, then $X_1 - X_2$ is not a Poisson variate.

Proof. Since X_1 and X_2 are independent variates, we have

$$M(t : X_1 - X_2) = M(t : X_1)\,.\,M(-t : X_2) = e^{\lambda_1(e^t - 1)} e^{\lambda_2(e^{-t} - 1)} = e^{\lambda_1(e^t - 1) + \lambda_2(e^{-t} - 1)}.$$

Since the right side cannot be expressed as $e^{\lambda(e^t - 1)}$, we infer that $Z = X_1 - X_2$ is not a Poisson variate. The proof also follows by trivial considerations : $X_1 \ge 0$, $X_2 \ge 0$, but $-\infty < Z < \infty$.

As $Z = X_1 - X_2$ can assume negative values, it cannot be a Poisson variate.

Aliter. Mean $(X_1 - X_2) \ne \text{Var}(X_1 - X_2) \Rightarrow X_1 - X_2$ is not Poissonian. $[\lambda_1 - \lambda_2 \ne \lambda_1 + \lambda_2]$

Example : Let $X \sim$ Pois (λ) and Y ~ Pois (μ). If $\lambda < \mu$, prove that

$P(X \le k) > P(Y \le k)$ [k : any integer].

Solution. Consider an auxiliary r.v. $Z \sim$ Pois ($\mu - \lambda$) which is independent of X. By reproductive property, $(X + Z) \sim$ Pois (μ) so that $X + Z$ has the same distribution as Y. Consequently

$$P(Y \le k) = P(X + Z \le k) = \sum_{r=0}^{k} P(Z = r)\,P(X \le k - r) \le P(X \le k) \sum_{r=0}^{k} P(Z = r)$$

$$= P(X \le k)\,.\,P(Z \le k) < P(X \le k), \text{ [since } P(Z \le k) < 1].$$

Thus $P(Y \le k) < P(X \le k)$. This result says that the smaller the parameter λ, the more the Pois (λ) is shifted to the left.

13-61. *Binomial approximation of Poisson's Distribution*

Suppose X is Pois (λ) and express X as a sum of n indep. Poisson variate X_i ; where each X_i is Pois (λ/n). Thus $X = X_1 + X_2 + \ldots + X_n$.

Now choose n in such a way that $P\{X_i > 1\}$ is negligible which implies that λ/n must be very small. Consequently the variate X_i is necessarily a Bernouli variate (zero-one variate), and so

$$P\{X_i = 0\} = e^{-\lambda/n} = 1 - (\lambda / n);\ P\{X_i = 1\} = 1 - P\{X_i = 0\} = \lambda / n.$$

With this interpretation, X can be viewed as bin n, p, where $p = \lambda/n$ and so

$$e^{-\lambda}\frac{\lambda^k}{k!} = P(X = k) = \binom{n}{k} p^k q^{n-k}.$$

This approximation holds good when $p = \lambda/n$ is small, n large and $\lambda = np$ 'moderate'

Remark. This proof builds binomial distribution from Poisson law.

13-62. *Worked-out Problems*

***Example* 1.** If X is Poisson with parameter λ, show that $(X - \lambda)/\sqrt{\lambda}$ is a standard variate. Find m.g.f. for this variate and show that it tends to $e^{t^2/2}$ as $\lambda \to \infty$.

Solution. See §11-31. Example 2*.

Important Note. For a $N(0, 1)$, $M(t : X) = e^{t^2/2}$. Hence what we have proved in essence is that : "As $n \to \infty$, p remains moderate, $np = \lambda \to \infty$, then Poisson distribution tends to normal distribution". Another proof of Normal approximation to the Poisson Distribution, using the density function and not m.g.f., is given in the chapter 17.

***Example* 2.** Let X and Y be independent Poisson variates. Show that the conditional distribution of X, given $X + Y = n$, is binomial.

Solution. Let $X \sim$ Pois (λ_1) and $Y \sim$ Pois (λ_2). Since X and Y are independent, by additive property, $Z = X + Y$ is Pois ($\lambda_1 + \lambda_2$). Now

$$P\{X = r \mid X + Y = n\} = \frac{P\{X = r, X + Y = n\}}{P(X + Y = n)} = \frac{P\{X = r, Y = n - r\}}{P(X + Y = n)} = \frac{P(X = r)\,P(Y = n - r)}{P(X + Y = n)}$$

$$= \frac{e^{-\lambda_1}(\lambda_1)^r}{r!}\,\frac{e^{-\lambda_2}(\lambda_2)^{n-r}}{(n-r)!} \Big/ \frac{e^{-\lambda_1 + \lambda_2}(\lambda_1 + \lambda_2)^n}{n!} = \frac{n!}{r!(n-r)!}\left(\frac{\lambda_1}{\lambda_1 + \lambda_2}\right)^r \left(\frac{\lambda_2}{\lambda_1 + \lambda_2}\right)^{n-r}$$

$$= \binom{n}{r} p^r q^{n-r}, \text{ where } p = \frac{\lambda_1}{\lambda_1 + \lambda_2}, q = 1 - p$$

Hence the conditional distribution of X, given $X + Y = n$ is bin $[n, \lambda_1 / (\lambda_1 + \lambda_2)]$.

***Example* 3.** If X and Y are independent Poisson variates with means λ_1 and λ_2, find the probability of (i) $X + Y = k$ (ii) $X = Y$.

Solution. Since X and Y are independent

$$G(t:X+Y) = x\,G(t:X)\,G(t:Y) = e^{\lambda_1(t-1)}.e^{\lambda_2(t-1)} = e^{(\lambda_1+\lambda_2)(t-1)}.$$

This shows that $Z = X + Y \sim$ Pois $(\lambda_1 + \lambda_2)$. hence

$$P\{Z=k\} = e^{-(\lambda_1+\lambda_2)}(\lambda_1+\lambda_2)^k / k!, \quad k=0,1,2,\ldots$$

(ii) Independence of X and Y is given by

$$G(t:X-Y) = G(t:X)\,G(t^{-1}:Y) = e^{\lambda_1(t-1)}e^{\lambda_2(t^{-1}-1)} = e^{-(\lambda_1+\lambda_2)+\lambda_1 t+\lambda_2 t^{-1}} \qquad \ldots(A)$$

$$P\{X-Y=0) = \text{Coeff. of } t^{\circ} \text{ in } G(t:X-Y) = \text{const. term in (A)}$$

$$= e^{-(\lambda_1+\lambda_2)} \times \text{const. term in} \left(\sum_{k=0}^{\infty}\frac{(\lambda_1 t)^k}{k!}\right)\left(\sum_{j=0}^{\infty}\frac{(\lambda_2 t^{-1})^j}{j!}\right)$$

$$= e^{-(\lambda_1+\lambda_2)}\,\Sigma(\lambda_1\lambda_2)^k/(k!)^2, \qquad k=0,1,2,\ldots [\text{Take } j=k]$$

Aliter.

$$P(X=Y) = \sum_{k=0}^{\infty} P(X=k, Y=k) = \sum_{k=0}^{\infty} P(X=k), (Y=k)$$

$$= \Sigma e^{-(\lambda_1+\lambda_2)}(\lambda_1\lambda_2)^k/(k!)^2 = e^{-(\lambda_1+\lambda_2)}\,\Sigma(\lambda_1\lambda_2)^k/(k!)^2,\ 0\le k<\infty.$$

***Example* 4.** Let $X \sim$ Pois (λ_1) and $Y \sim$ Pois (λ_2) be independent and let $Z = X - Y$. Prove that

$$P\{Z=r\} = \text{coefficient of } t^r \text{ in } \exp\{-\lambda_1-\lambda_2+\lambda_1 t+\lambda_2 t^{-1}\}.$$

Deduce the p.m.f. of Z when $\lambda_1 = \lambda_2$.

Solution. Since X and Y are indep. the p.g.f. yields

$$G(t:X-Y) = G(t:X).G(t:-Y) = G(t:X)\,G(t^{-1}:Y)$$

$$\therefore \quad G(t:Z) = e^{-\lambda_1(t-1)}.e^{\lambda_2(t^{-1}-1)} = e^{-\lambda_1-\lambda_2+\lambda_1 t+\lambda_2 t^{-1}}$$

So,

$$p_r = P\{Z=r\} = \text{coeff. of } t^r \text{ in } G(t:Z) \qquad \text{[Def. of p.g.f.]}$$

$$P_r = \text{Coeff. of } t^r \text{ in exp } \{-\lambda_1-\lambda_2+\lambda_1 t+\lambda_2 t^{-1}\}. \qquad \ldots(1)$$

Deduction. Using exponential series, we can rewrite (1) as

$$p_r = e^{-(\lambda_1+\lambda_2)} \text{ coeff. of } t^r \text{ in} \left(\sum_{j=0}^{\infty}\frac{(\lambda_1 t)^j}{j!}\right)\left(\sum_{k=0}^{\infty}\frac{(\lambda_2 t^{-1})^k}{k!}\right)$$

$$= e^{-(\lambda_1+\lambda_2)}\sum_{k=0}^{\infty}\frac{(\lambda_1)^{k+r}(\lambda_2)^k}{k!(k+r)!}. \qquad [\text{using } j-k=r]$$

When $\lambda_1 = \lambda_2 = \lambda$ (say), then

$$p_r = e^{-2\lambda}\left(\sum_{k=0}^{\infty} \frac{\lambda^{2k+r}}{k!(k+r0!}\right) = e^{-2\lambda}\, I_r(2\lambda) \qquad \ldots(2)$$

where $I_r\,(2\lambda)$ is Bessel's modified function of order r and argument 2λ.

Note. We can obtain (2) directly as under

$$G(t:Z) = G(t:X)\,G(t:-Y) = G(t:X)\,G(t^{-1}:Y) = e^{\lambda(t-1)}\,.\,e^{\lambda(t^{-1}-1)} = e^{-2\lambda+\lambda(t+t^{-1})}$$

$$p_r = P(Z=r) = \text{Coeff. of } t^r \text{ in } G(t:Z) = e^{-2\lambda} \text{ Coeff. of } t^r \text{ in}$$

$$\left(\sum_{j=0}^{\infty} \frac{(\lambda t)^j}{j!}\right)\left(\sum_{k=0}^{\infty} \frac{(\lambda t^{-1})^k}{k!}\right) = e^{-2m} \sum_{k=0}^{\infty} \frac{(\lambda^{2k+r})}{k!(k+r)!} \qquad [j-k=r].$$

Problems with Solutions Provided at the End of the Text

1*. Is the Poisson distribution skewed to the right or left ? Compute Pearson's measure of skewness for this distribution.

2*. Deduce the first four moments about the mean of the Poisson distribution from those of the Binomial distribution.

3*. Let $X_1, X_2, \ldots, X_n$ be i.i.d Pois (λ) variates. Find the distribution of the statistic $\overline{X}$.

4*. X_1 and X_2 are two independent, variates X_1 is Pois (λ) and $X_1 + X_2$ is Pois ($\lambda + \mu$). Find the distribution of X_2.

5*. If $X \sim$ Pois (N) and r.v. N possesses m.g.f. $M_N(t)$, find $M_X(t)$.

6*. Let $X = X_1 + X_2 + \ldots + X_N$, where X_j ($j = 1, 2, \ldots$) are independent and N is Pois (θ). Find $M_X(t)$. Simplify the result if X_j are i.i.d.

7*. Suppose that $M(t : X) = e^{\lambda(e^t - 1)}, x = 0, 1, 2, \ldots$. Find the p.m.f. of X.

8*. The m.g.f. of a variate X is $e^{4(e^t - 1)}$. Show that

$$P\{\mu - 2\sigma < X < \mu + 2\sigma\} = 0.931.$$

9*. Show how the m.g.f. of binomial distribution tends to that of the Poisson distribution.

10*. The number of deaths per day in a city due to road accidents and due to other causes follow Poisson distribution with parameters 2 and 6 respectively. Find the probability that the total number of deaths in a day is 10 or fewer. What is the maximum probable number of deaths in a day.

11*. If X_i, $1 \le i \le k$ are indep. Poisson variates, show that the conditional distribution of X_1, given $X_1 + \ldots + X_k$ is binomial.

Exercise 13(c)

1. If X is Pois (λ), show that $E(X^2) = \lambda\, E(X + 1)$.
2. (a) Identify the distribution with $M_X(t) = \exp 3(e^t - 1)$. Find $E(X)$ also.

   (b) Identify the variates X and Y when $M_X(t) = (1/27)(1 + 2e^t)^3 \cdot e^{3(e^t - 1)}$,

   $M_Y(t) = (1/32)(1 + e^t)^5 \cdot e^{-2(1-e^t)}$.
3. (a) If X is a Poisson variate and $f(0) = f(1)$ show that $E(X) = 1$.

   (b) If X is a Poisson variate and $f(0) = 1/2$, show that Var $(X) = \ln_e 2$.

   (c) Let $X \sim$ Pois (λ). Show that $E(e^{-kX}) = \exp[-\lambda(1 - e^{-k})$ and find $E(Xe^{-kX})$. Hence show that, if X is the mean of n indep. variate $X_1, \ldots, X_n$ each Pois (λ) distributed then exp $(-\overline{X})$ as an estimate of exp $(-\lambda)$ is biased, although $\overline{X}$ is an unbiased estimate of λ.
4. If for a Poisson variate X, $f(1) = 2f(2)$, show that $f(0) = e^{-1}$ and $\sigma_x^2 = 1$.
5. Let X be a Poisson variate such that $E(X^2) = 6$. Show that $E(X) = 2$.
6. If $X \sim$ Pois (λ) and $Y = X^2$, show that Corr $(X, Y) = (2\lambda + 1)/\sqrt{(1 + 6\lambda + 4\lambda^2)}$.
7. Let $X \sim$ Pois (λ). Find the m.g.f. of $Y = 2X - 1$ and deduce that Var $(Y) = 4\lambda$.
8. If X is Pois (6), show that $P(0 < X < 9) > 1/3$.
9. The number of aeroplanes arriving at an airport in a 30 minute-interval obeys the Pois (25). Use Chebyshev's inequality to show that the least chance that the number of planes to arrive within a 30 minute interval will be between 15 and 35 is 0.75.
10. Suppose that the number of motor cars arriving in a certain parking lot in any 15 minutes period obeys the Pois (80) law. Use Chebyshev's inequality to show that a lower bound for the probability that the number of motor cars arriving in a given 15 minute period will be between 60 and 100 is 0.80.
11. A toll tax of Rs. 5/- is collected by the Government from the drivers of every car and bus that crosses a certain bridge. The number/s of cars and buses that cross the bridge in a day are random variables having Poisson distribution with mean 5 and 4 respectively. Show that the probability that no toll tax will be collected on a certain day is e^{-9}.
12. Let X be a Poisson variate such that $9f(4) + 90 f(6) = f(2)$. Show that (i) $E_X = 1$, (ii) g_1, the coefficient of skewness = 1.
13. (a) Let X and Y be indep. Poisson variates with $\sigma_x^2 = 4\sigma_y^2$. Find $P(X + Y \geq 2)$.

    (b) Let X and Y be two indep. Poisson r.v.s with means 1 and 2 respectively. Show that

    (i) $P(X + Y \geq 4) = p$? (ii) $P\{X + Y < 4\} = \Sigma(e^{-3} 3^x / x!), 0 \leq x \leq 3$.
14. If X and Y are independent Poisson variates such that $P(X = 1) = P(X = 2)$, $P(Y = 2) = P(Y = 3)$. Find $P(X = 4)$. Show that Var $(X - 2Y) = 14$.
15. If X and Y are independent Poisson variates with mean 1 and 3 respectively, find the m.g.f. of $Z = 3X + Y$, and show Var $(Z) = 12$.
16. Prove that the function that generates the central moments for Pois (λ) is $M(t) = \exp\{\lambda(e^t - t - 1)\}$. Show that this function satisfies the differential equation $dM/dt = \lambda t\, M(t) + \lambda\, dM / d\lambda$.

17. If X is Pois (λ) and μ_r the rth central moment, then

$$\mu_{r+1} = \lambda \sum_{j=1}^{r} \binom{r}{j} \mu_{r-j} = \lambda\left[\binom{r}{1}\mu_{\mu-1} + \binom{r}{2}\mu_{r-2} + \ldots + \binom{r}{r}\mu_0\right]$$

18. If X is Pois (λ), and $\mu_r' = \mathsf{E}(X)^r$ show that

$$\mu_{r+1}' = \lambda[\mu_r' + (d\mu_r' / d\lambda)].$$

19. If X is Pois (λ), show that
(i) $P(X \le \lambda/2) \le 4/(\lambda+4) < 4/\lambda$; (ii) $P(X \ge \lambda) < 1/(\lambda+1) < 1/\lambda$.

20. Write $X = \mu[1 + (X-\mu)/\mu]$, where $\mu = \mathsf{E}(X)$ and $\sigma^2 = \text{Var}(X)$. Show that $\mathsf{E}(\sqrt{X}) \sim \sqrt{\mu}\,[1 - (\sigma^2/8\mu^2)]$ and if X is Poisson distributed, then Var $(\sqrt{X})$ is approximately 1/4.

21. Let $X \sim$ Pois (λ) and $Y \sim$ Pois (μ) be independent. Find the joint p.g.f. of X and $Z = X + Y$. Deduce that

$$P(X = r \mid Z = n) = \binom{n}{r} q^{n-r} p^r, \left(p = \frac{\lambda}{\lambda+\mu}, q = 1-p\right), \quad 0 \le x \le n.$$

13-70. *Miscellaneous Worked-Out Problems*

***Example* 1.** If $X \sim$ Pois (λ), find E (cos tX) and E (sin tX).

Solution. By Euler's Thm : $e^{itX} = \cos tX + i \sin tX$. Hence

$$\mathsf{E}(\cos tX + i \sin tX) = \mathsf{E}(e^{itX}) \qquad \text{[Use m.g.f. of Pois } (\lambda)]$$

$$= e^{\lambda(e^{it}-1)} = e^{\lambda[(\cos t - 1) + i \sin t]}$$

$$= e^{\lambda(\cos t - 1)}\,[\cos(\lambda \sin t) + i \sin(\lambda \sin t)]$$

Equating real and imaginary parts.

$$\mathsf{E}(\cos tX) = e^{\lambda(\cos t-1)} \cos(\lambda \sin t), \quad \mathsf{E}(\sin tX) = e^{\lambda(\cos t - 1)} \sin(\lambda \sin t).$$

***Example* 2.** If X and Y are correlated variates each having Poisson distribution, show that $X + Y$ cannot be Poisson-distributed.

Solution. If $X \sim$ Pois (λ) and $Y \sim$ Pois (μ), then $\mathsf{E}(X) = \lambda = \text{Var}(X)$, $\mathsf{E}(Y) = \mu = \text{Var}(Y)$
Since X and Y are correlated, so $\rho = \text{Corr}(X, Y) \ne 0$

$$\mathsf{E}(X+Y) = \mathsf{E}(X) + \mathsf{E}(Y) = \lambda + \mu$$

$$\text{Var}(X+Y) = \text{Var}(X) + \text{Var}(Y) + 2\,\text{Cov}(X, Y) = \lambda + \mu + 2\sqrt{\lambda\mu}\,\rho, (\rho \ne 0)$$

Since $\mathsf{E}(X + Y) \ne \text{Var}(X + Y)$, it follows that $X + Y$ is not Poisson-distributed.

Note. Since X and Y are related, let $Y = kX$ (or $Y = \varphi(X)$]. Then $Z = X + Y = (k + 1) X$; and Z may assume even fractional values; so that Z is not a Poisson r.v.

***Example* 3.** Let $X \sim$ Pois (λ_1) and $Y \sim$ Pois (λ_2) be independent variates. Find $\mathsf{E}\{X \mid (X+Y=n)\}$.

Solution. By addition theorem, $(X+Y) \sim$ Pois $(\lambda_1+\lambda_2)$.

$$\therefore \quad P\{(X+Y)=n\} = \frac{e^{-(\lambda_1+\lambda_2)}(\lambda_1+\lambda_2)^n}{n!} \qquad \text{...(i)}$$

Further $P\{X=k,\, X+Y=n\} = P\{X=k,\, Y=n-k\}$.

$$= P(X=k)\,.\,P(Y=n-k), \qquad \text{[By Indep.]}$$

$$= \frac{e^{-\lambda_1}(\lambda_1^k)}{k!}\cdot\frac{e^{-\lambda_2}(\lambda_2)^{n-k}}{(n-k)!} = \binom{n}{k}\frac{e^{-(\lambda_1+\lambda_2)}.(\lambda_1)^k(\lambda_2)^{n-k}}{n!} \qquad \text{(ii)}$$

$$\therefore \quad P\{X=k \mid (X+Y)=n\} = \frac{P\{X=k,(X+Y)=n\}}{P\{X+Y)=n\}}$$

$$= \binom{n}{k}\left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^k\left(\frac{\lambda_2}{\lambda_1+\lambda_2}\right)^{n-k} = \binom{n}{k}q^{n-k}\,p^k$$

where $p = \lambda_1/(\lambda_1+\lambda_2)$. Thus the conditional probab. is bin (n, p) and its mean is np. So

$$\mathsf{E}[X \mid (X+Y=n)] = np = n\lambda_1/(\lambda_1+\lambda_2).$$

***Example* 4.** If X is bin (N, p) where N itself is Pois (λ), find the characteristic function of X and hence deduce that X is Pois (λp).

Solution. $\quad \phi(t:X) = \mathsf{E}(e^{itX}) = \mathsf{E}_N\mathsf{E}\,e^{itX}(N) = \mathsf{E}(q+pe^{it})^N. \qquad$ [Double-E Rule]

$$= \mathsf{E}(\theta^N) = e^{\lambda(\theta-1)} = e^{\lambda p(e^{it}-1)} \qquad \text{[p.g.f. of Pois } (\lambda)]$$

This shows that $X \sim$ Pois (λp).

***Example* 5.** The joint density of variates X and Y is

$$f(x, y) = e^{-2}/[x!(y-x)!], \quad y=0,1,2,\ldots;\; x=0,1,\ldots,y$$

Find the m.g.f. $M(t_1, t_2)$, Corr (X, Y) and the marginal distributions.

Solution. With the usual standard notation

$$M(t_1,t_2) = \mathsf{E}\{e^{t_1X+t_2Y}\} = \sum_{y=0}^{\infty}\sum_{x=0}^{y} e^{t_1x+t_2y}\frac{e^{-2}}{x!(y-x)!} = e^{-2}\sum_{y=0}^{\infty}\frac{e^{t_2y}}{y!}\sum_{x=0}^{y}\binom{y}{x}e^{t_1x}$$

$$= e^{-2}\sum_{y=0}^{\infty}\frac{e^{t_2y}}{y!}(1+e^{t_1})^y = e^{-2}\sum_{y=0}^{\infty}\frac{[e^{t_2}(1+e^{t_1})]^y}{y!} = e^{-2}.\exp(1+e^{t_1})e^{t_2}$$

$$= \exp[-2+(1+e^{t_1})e^{t_2}]$$

From this we get immediately

$$M(t_1,0) = \exp(e^{t_1}-1) \Rightarrow X \sim \text{Pois}(1).\ M(0,t_2) = \exp 2(e^{t_2}-1) \Rightarrow Y \sim \text{Pois}(2).$$

Thus the marginal distribution are Poisson variates with parameters 1 and 2 respectively.

And as such $\mathsf{E}(X) = \sigma_X^2 = 1; 2 = \mathsf{E}(Y)\,\sigma_Y^2$.

Now $\qquad K(t_1,t_2) = \ln M(t_1,t_2) = e^{t_2}(1+e^{t_1}) - 2$

$\therefore \qquad \partial^2 K/\partial t_1\,\partial t_2 = e^{t_1+t_2},\ \sigma_{XY} = \partial^2 K(0,0)/\partial t_1\,\partial t_2 = 1; \rho(X,Y) = \sigma_{XY}/\sigma_X\sigma_Y = 1/\sqrt{2}$

Remark. $M(t_1, 0)\,.\,M(0, t_2) \neq M(t_1, t_2) \Rightarrow X$ and Y are dependent.

Problems with Solutions Provided at the End of the Text

1*. Let $Y_j \sim$ bin $(1, p)$ be a sequence of i.i.d. variates; and let $X \sim$ Pois (λ) be indep. of Y's. Show that $S = Y_1 + Y_2 + \ldots + Y_X$ and $D = X - S$ are independently distributed.

2*. Let $X \sim$ Pois (λ), and $Y = X_k$ (k is real number). Find the mean and variance of Y. Deduce that

$$\mathsf{E}(\sqrt{X}) = \sqrt{\lambda}\left\{1 - \frac{1}{8\lambda} - \frac{7}{128\lambda^2} + \ldots\right\},\ \text{Var}(\sqrt{X}) = \frac{1}{4} + \frac{3}{32\lambda} + \ldots$$

3*. The number X of blackflies on a board bean leaf is assumed to be a Pois (λ). A plant inspector, however, records the number on a leaf only if there are blackflies present. What is the probability that he records x blackflies on one leaf ? What is the expected number of blackflies recorded per leaf ? What is the probability that the total number of blackflies from two recorded leaves is n ?

4*. Let X be any non-negative integer-valued variate and a be any positive number. Show that $P\{X \le a\} \le t^{-a}\,\mathsf{E}(t^X),\quad 0 \le t \le 1.$ ...(1)
Verify the inequality : $P\{X \le \lambda/2\} \le (2/e)^{(\lambda/2)}$, when X is Pois (λ).

5*. Let X and Y be i.i.d Pois (1) variates. If $Z = \min\{X, Y\}$, find $P(Z = 1)$.

6*. Find Corr (X, Y), when a bivariate distribution of (X, Y) has the p.g.f.

$$G(s, t) = \exp[a(s-1) + b(t-1) + c(s-1)(t-1)], \qquad (a, b, c \text{ all positive})$$

Miscellaneous Exercises

1. Let X and Y be independent Poisson variates with Var $(X + Y) = 9$ and $P\{X = 3 \mid X + Y = 6\} = 5/54$. Show that $\mathsf{E}(X) = (9 \pm 3\sqrt{3})/2$.
2. Let X by Pois (λ). The conditional distribution of $Y - 1$, given $X = x$, is Poisson with $\mathsf{E}(Y-1) = mx$. Show that for $y = 1, 2, \ldots$

$$P[Y = y] = \sum_{j=0}^{\infty}\left[\frac{\lambda^j e^{-\lambda}}{\lambda!}\cdot\frac{(jm)^{y-1}}{(y-1)!}e^{-jm}\right] \text{ and } \mathsf{E}(Y) = 1 + \lambda m,\ \text{Var}(Y) = \lambda m(1+m).$$

3. If X is Pois (λ), show that $\mathsf{E}(t^X \mid X > 0) = (e^{\lambda t} - 1)/(e^{\lambda} - 1)$ and hence find $\mathsf{E}(X \mid X > 0)$. and Var $(X \mid X > 0)$.

4. A truncated Poisson distribution is given by mass function

$$f(x) = e^{-\lambda}\lambda^x / x!(1-e^{-\lambda}); \qquad x = 1, 2, ...$$

Find the m.g.f. and $E(X^{(r)})$. Hence or otherwise find the mean and the variance of the distribution.

5. Variates X_i, $i = 0, 1, 2, ...$ are independent and $E(X_i) = i$. N is Pois (λ) and is independent of the X_i show that $E(X_1 + X_2 + ... + X_n) = \lambda(2+\lambda)/2$.

6. For $X \sim$ Pois (λ), verify the following inequalities :

(a) $P(X \le \frac{1}{2}\lambda) < (4/\lambda)$. (b) $P(X \ge 2\lambda) \le (1/\lambda)$.

7. If $X \sim$ Pois (λ), show that $E(1+X)^{-1} = (1-e^{-\lambda})/\lambda$ and $E[(1+X)(2+X)]^{-1} = [1-(1+\lambda)e^{-\lambda}]/\lambda^2$.

8. If X is any non-negative integer-valued variate and a is any positive number, show that $P(X \ge a) \le t^{-a} E(t^X)$; $0 < t < h$; $P(X \le a) \le t^{-a} E(tX), -h < t < 0$.
Verify the inequality : $P(X \ge 2\lambda) \le (e/4)^{\lambda}$, when X is Pois (λ).

9. (a) A catepillar $2a$ inches long starts to cross at rt angles a one-way cycle track H yards wide, at a speed of f ft/sec. If cycles are passing this particular spot at random intervals but at an average rate of λ per second, what is the probability that the caterpillar reaches the other side safely ? Assume that the impress of a cycles tyre on the ground equals h inches and that the caterpillar may only be touched to be considered hurt.

(b) At a busy street inter-section, it is estimated that a joy-walker will be hit by a car with probability 0.01. Assuming individual trips form independent trials, find the chance that the joywalker remains unhit if he crosses the street twice per day for 50 days.

10. The number n of carriers of a disease in a certain district is Pois (λ). The number of persons infected by different carriers are independent. Poisson variates, each with parameter θ. Find the p.g.f. for the total number of persons infected. Hence obtain the arthmetic mean and variance of this distribution.

11. In a species of animals for which the sex-ratio at birth is 1 : 1, the probability of a litter being of size r is proportional to the Poisson probability of r with mean λ. Prove that the prob. that all the animals in any one litter are of the same sex is $2/(1 + e^{\lambda/2})$.

12. A radio-active source emit on the average 2.5 particles per second. Show that the probability that 3 or more particles will be emitted in an intervals of 4 seconds is $\Sigma e^{-10}(10)^x / x!$, $3 \le x < \infty$.

13. (a) The probability that a radio-active substance gives off n beta-particles in a unit of time is Poisson (λ), $n = 0, 1, 2, ...$ The probability that a given particle will strike a counter and be registered is p. Show that the probability of registering n beta-particles in a unit of time is also Poisson. Find also the mean number of particles registered in a unit time.

(b) The number of female insects in a given region follows Pois (λ) distribution. The number of eggs laid by each female insect follows Pois (λ) distribution. Find the probability of the number of eggs in the region and its variance.

14. It is known that if the average number of bacteria per litre of suspension in a reservoir is 100 λ, the number of bacteria present 1 c.c. specimen of liquid is a Poisson (λ). Find the probability that, of five independent specimens, each of size 1 c.c. exactly 1 will contain no bacteria. Show that this probability reaches a maximum when $\lambda = \ln 5$.

15. The suicide rate in state A is 5 suicides per one million persons per month. Find the chance that in a certain city of population 500,000 there will be at most 5 suicides in a month, what if during 1 year there were atleast 2 months in which more than 5 suicides occurred ?

16. In a certain experimental procedure, a given quantity of radio-active matter is available, which emits $K(t)$ particle in t seconds, where $K(t)$ is Pois $(10^5 t)$. It is required to choose an 'exposure time' so that (a) on the average the number of particles emitted during the exposure is not less than 10^2, and more 10^4. (b) The coefficient of variation of particles emitted during the exposure does no exceed 3.2 %.

17. In 1000 consecutive issues of the *Utopian Seven Daily Chronicle*, the deaths of centenarians were recorded, the number x having frequency f according to the table

x :	0	1	2	3	4	5	6	7	8
y :	229	325	257	119	50	17	2	1	0

Show that the distribution is roughly Poissonion by calculating its mean and then the frequencies in the Poissonian distribution with the same mean and the same total frequency of 1000. Find also the variance of the given distribution. ($e^{-1.5} = 0.2231$)

18. Letters were received in an office on each of 100 days. Assuming the following data to form a random sample from a Poisson distribution, find the expected frequencies correct to the nearest unit taking $e^{-4} = 0.0183$.

No. of letters :	0	1	2	3	4	5	6	7	8	9	10
Frequency :	1	4	15	22	21	20	8	6	2	0	1

[**Ans.** 2, 7, 15, 20, 16, 10, 6, 3, 1, 0]

19. From records of 10 Russian Army Corps kept over 20 years, the following data was obtained showing the number x of deaths caused by the kicks of a horse. Determine m, the average number of deaths per Army Corps per annum, and calculate the theoretical Poisson Frequencies and compare them with the observed frequencies of occurrences f

x :	0	1	2	3	4	Total
f :	109	65	22	3	1	200

[**Ans.** $e^{-0.61} = 0.5436$]

Love is photogenic, it needs darkness to develop.

A lie told often becomes the truth. (Lenin)

Geometric Distribution. Negative Binomial Distribution 14

14-10. *Geometric Distribution*

Definition. A random variable X possessing the following probability law

$$P(X = x) = f(x) = q^x p, \quad x = 0, 1, 2, \ldots, f(x) = 0, \text{ otherwise} \qquad \ldots(1)$$

is called a geometric variate. The distribution (1) is called the geometric distribution. We often express (1) by writing 'X' as geom (p) or $X \sim$ geom (p).

Notes : **1.** Geometric distribution is a particular (but most important) case of Neg-Bin distribution, when $k = 1$. Thus Neg-bin $(1, p)$ = geom (p). [See § 14-60]

2. From (1), $\sum_{x=0}^{\infty} pq^x = p \sum_{x=0}^{\infty} q^x = \frac{p}{1-q} = \frac{p}{p} = 1; \quad f(x) = 0, \text{otherwise}$

It follows that (1) is a legitimate probability mass function.

3. As the various terms in (1) are the terms of a G.P.; the distribution has come to be known as 'geometric distribution'.

Physical Conditions for the Occurrence of Geometric Distribution

1. The result of each trial (experiment) can be classified into one of two disjoint categries, say, success (S) and failure (F).

2. P (Success) $= p$, remains the *same constant* for each trial.

3. The outcomes of all trials are independent of each other.

4. The series of trials is performed a *variable* number of times until the first success is achieved.

Note. Only the 4th condition differs from the conditions of Binomial distribution.

14-20. *Two Forms of Geometric Distribution : geom (p) and gem (p)*

Form **1.** Let X denote the number of failures preceding the *first* success in a sequence of independent Bernoulli's trials. Then $P\{X = x\} = q^x p$, $x = 0, 1, 2, 3, \ldots$

Form **2.** Let Y denote the number of trials required to obtain the first success in a sequence of independent Bernoulli's trials. Then $P\{Y = y\} = q^{y-1} p$, $y = 1, 2, 3, \ldots$.

Relation. $Y = X + 1$. [$E(Y) = 1 + E(X)$, $\quad \text{Var}(Y) = \text{Var}(X)$]

Since $x + 1$ trials are needed to have x failures and one success, we must have $y = x + 1$.

Modelling. **(1)** We define the events as under :

A = {In the first x trials, there is no (i.e. zero) success}, B = {$(x + 1)$th trial is the first success}

Then, $P(A) = {}^xC_0\, q^x p^0 = q^x$. [Bin. p.m.f.] $\quad P(B) = p; \quad P(AB) = P(A)\,P(B) = q^x p.$

$\therefore \quad P(X = x) = q^x p, x = 0, 1, 2, 3, \ldots$

Modelling. (2) We define the events as under :

A_1 = {In the first $(y-1)$ trials, there is zero (i.e. no success} ; B_1 = {yth trial is the first success}.

Then $P(A_1) = {}^{y-1}C_0\, q^{y-1} p^0 = q^{y-1}$ [Binomial p.d.f.]; $P(B_1) = p$, $P(A_1 B_1) = P(A_1)\, P(B_1) = q^{y-1} p$

$$\therefore \quad P\{Y = y\} = q^{y-1}\, p, \; y = 1, 2, 3, \ldots$$

Notation. For the *support*, $x = 0, 1, 2, \ldots, f(x) = q^x p$, we write $X \sim$ geom (p).
For the *support* $x = 1, 2, 3, \ldots, f(x) = q^{x-1} p$, we write $X \sim$ geom (p). It is often called *Pascal* geometric distribution.

Comments. Suppose trials are continued until the event (S) occurs for the first time. The number X of trials performed **before** the event S occurs, is geom (p).

If the number of trials counted (Y) includes the trial in which the event S occurs **for the first time** (written f.f.t.) then Y is gem (p). As such it is also notated $Y \sim$ fft(p).

14-21. *Survival Functions and c.d.f.*

Let $X \sim$ geom (p) and $Y \sim$ geom (p); then, survival function is $P(Z > c)$. Now

$$P(X > a) = \sum_{a+1}^{\infty} pq^x = pq^{a+1}(1 + q + q^2 + \ldots) = \frac{pq^{a+1}}{1-q} = q^{a+1}.$$

$$P(X \geq a) = \sum_{a}^{\infty} pq^x = pq^a(1 + q + q^2 + \ldots) = \frac{pq^a}{1-q} = q^a.$$

$$F_X(a) = P(X \leq a) = 1 - P(X > a) = 1 - q^{a+1}. \text{ [c.d.f.]}.$$

$$P(Y > b) = \sum_{b+1}^{\infty} pq^{y-1} = pq^b(1 + q + q^2 + \ldots) = \frac{pq^b}{1-q} = q^b.$$

$$P(Y \geq b) = \sum_{b}^{\infty} pq^{y-1} = pq^{b-1}(1 + q + q^2 + \ldots) = \frac{pq^{b-1}}{1-q} = q^{b-1}.$$

$$F_Y(b) = P(Y \leq b) = 1 - P(Y > b) = 1 - q^b.$$

14-22. *Mean, Variance, Median and Modal Value*

For $X \sim$ geom (p) ; $\quad E(X) = \Sigma\, x\, pq^x$; $\quad E(X^2) = \Sigma x^2\, pq^x, \;\; 0 \leq x < \infty.$ ...(1)

Differentiating the geometric series $\Sigma q^x = q/(1-q), \;\; 1 \leq x < \infty$ we obtain

$$\Sigma x q^{x-1} = 1/(1-q)^2 \Rightarrow \Sigma x q^x = q/(1-q)^2. \quad (1 \leq x < \infty) \qquad \ldots(2)$$

Substituting (2) into 1(a) gives $E(X) = p.q/(1-q)^2 = q/p$.

Differentiating the series (2) we obtain

$$\Sigma x^2 q^{x-1} = (1+q)/(1-q)^3, \text{ i.e. } \Sigma x^2 q^x = q(1+q)/(1-q)^3, \;\; 1 \leq x < \infty \qquad \ldots(3)$$

Substituting (3) in 1(b) gives $E(X^2) = pq\,(1+q)/(1-q)^3 = q(1+q)/p^2$.

$$\therefore \quad \text{Var}(X) = E(X^2) - E^2(X) = [(q+q^2)/p^2] - (q^2/p^2) = q/p^2.$$

Median. Let m = Med (X). By definition :

$$\frac{1}{2} = \sum_{x=0}^{[m]} pq^x = p(1+q+\ldots+q^{[m]}) = \frac{p(1-q^{m+1})}{1-q} = 1-q^{m+1}.$$

$\therefore \quad (m+1)\ \ln q = -\ln 2 \Rightarrow m = -1 - (\ln 2/\ln q).$

Note. Med (X) < Mean (X).

Modal Value. Since $f(x) = q^x\, p$, $x = 0, 1, 2, \ldots$ has the max. value only at $x = 0$, it follows that $X = 0$ is the mode of goem (p).

Y ~ gem (p). $Y = 1 + X \Rightarrow \mathrm{E}(Y) = 1 + \mathrm{E}(X) = 1 + (q/p) = (1/p)$. Var ($Y$) = Var ($X$)

Med Y = 1 + med $X = -(\ln 2/\ln q)$.

Modal value is at $X = 1$, since $f(x) = pq^{x-1}, x = 1, 2, 3, \ldots$ attains maximum only at $x = 1$.

Probability Recurrence Formula

$$f(x) = q^x p,\ f(x+1) = q^{x+1}\, p = q(q^x\, p) = q\, f(x).$$

Thus $\quad f(x+1) = q \,.\, f(x).$

14-23. *Worked-out Problems*

Example 1. A die is cast until a 6 appears. What is the chance that it must be cast more than five times ?

Solution. Here $p = 1/6$ (prob. of getting a six. If X is the number of tosses required for the first success ; then $P(X = k) = q^{k-1}\, p$, $k = 1, 2, \ldots$ so that

$$P(X > r) = P(X \geq r+1) = \sum_{k=r+1}^{\infty} P(X = k) = \sum_{r+1}^{\infty} (p\,q^{k-1}) = q^r$$

$$\therefore \quad P(X > 5) = (5/6)^5.$$

Example 2. Let X ~ geom (p) and Y ~ geom (p) be independent. Show that conditional distribution of $X \mid (X + Y) = n$ is uniform.

Solution. Here $\quad f(x) = pq^x,\ x = 0, 1, 2, 3, \ldots$ Now

$$M(t : X + Y) = [M(t : X)]^2 = [p/(1 - qe^t)]^2$$

$$\therefore \quad Z = (X + Y) \sim NB(2, p) \Rightarrow f(z) = (1 + z)\, p^2\, q^z \text{ i.e. } P(Z = n) = (1 + n)\, p^2 q^n$$

$$\therefore P\{X = r \mid X + Y = n\} = \frac{P\{X = r, Y = n - r\}}{P(X + Y = n)} = \frac{pq^r \,.\, pq^{n-r}}{(n+1)\, p^2\, q^n} = \frac{1}{n+1},\quad 0 \leq r \leq n.$$

Thus, conditional distribution is uniform over $S = \{0, 1, \ldots, n\}$.

Note. $f(z)$ can also be had through Multistage p-rule.

Example 3. Let X ~ geom (p) and Y ~ geom (p') be independent. Find p.m.f. of $Z = \min\{X, Y\}$. [§14-21]

Solution.

$$P\{Z > z\} = P\{X > z, Y > z\} = P(X > z)\,.\,P(Y > z) = (q^{z+1})\,(q')^{z+1}$$

$$\therefore \quad P\{Z = z\} = P\{Z \geq z\} - P\{Z \geq z+1\} = P\{Z > z - 1\} - P\{Z > z\}$$

$$= (qq')^z - (qq')^{z+1} = (1 - qq')\,(qq')^z$$

$$= Q^z P,\quad z = 0, 1, 2, \ldots \quad Q = qq',\quad P = 1 - Q$$

Thus, Z ~ geom (P).

Problems with Solutions Provided at the End of the Text

1*. If T_1 is the waiting time (number of tosses) for a coin C_1 to show a head and T_2 is the corresponding waiting time for a coin C_2 to produce a head, find $P\{T_2 > T_1\}$ where $P\{\text{coin } C_j \text{ shows head}\} = p_j$, $j = 1, 2$. The outcomes of C_1 and C_2 are independent.

2*. A and B shoot independently until each has hit his own target. They have probabilities 3/5, 5/7 of hitting the targets at each shot respectively. Find the probability that B will require more shots than A.

3*. A sample of 100 items is taken every hour from the output of a machine. If the sample contains 3 or more defective items, the machine is stopped for adjustment. If the machine is producing 3% defective items, find the average interval of time between successive adjustments.

4*. For the geometric p.m.f. $f(x) = 2^{-x} = 1, 2, 3, \ldots$ show that Chebyshev's inequality gives $P\{\,|X - 2| \le 2\} > \frac{1}{2}$ while the actual probability is 15/16.

5*. X and Y are i.i.d. geom (p) variates and $Z = \max\{X, Y\}$. Show that

$$P\{Z=m\} = 2pq^{m-1} - p(1+q)q^{2m-2},\quad (q=1-p),\quad 1 \le m < \infty.$$

14-30. Memoryless Property (or Markov Property) of Geometric Distribution

If X is geom (p), then X lacks memory in the sense that

$$P\{X > s+t \mid X > s\} = P\{X > t\}.$$

Proof. For geom (p), survival function : $P(X > a) = q^a$, $a = 0, 1, 2, \ldots$.
Therefore, $P\{X > s + t\} = q^{s+1}$, $P\{X > s\} = q^s$.

$$\text{So}: P\{X > s+t \mid X > s\} = \frac{P(X > s+t)}{P(X > s)} = \frac{q^{s+t}}{q^s} = q^t = P\{X > t\}.$$

Converse. Let s and t be integers. If the positive integral-valued variate X is such that

$$P\{X > s+t \mid X > s\} = P\{X > t\}, \text{ then } X \sim \text{geom}(p). \qquad \ldots(2)$$

Proof. Let $P(X > 1) = c$. Set $P(X = k) = p_k$ and $P(X > k) = g(k)$.

$$p_{s+1} = g(s) - g(s+1) \Rightarrow p_{s+1}/g(s) = 1 - [g(s+1)/g(s)] = 1 - P\{X > s+1 \mid x > s\} = 1 - g(1)$$

Thus $p_{s+1} = (1-c)\,g(s)$. $[c = g(1)]$. We use this result recursively : ...(3)

$$p_{s+1} = (1-c)\,g(s) = (1-c)[g(s-1) - p_s] = (1-c)[g(s-1) - (1-c)\,g(s-1)] = (1-c)\,cg(s-1).\,[\text{by } (2)]$$

$$= (1-c)\,c^2 g(s-2) = (1-c)\,c^3 g(s-3) = \ldots = (1-c)\,c^{s-1} g(1) = (1-c)\,c^s.$$

If we set $1 - c = p$, then $P(X = s) = pq^{s-1}$, $s = 1, 2, 3, \ldots$, i.e. $X \sim$ geom (p).

Remark. This property characterizes the geometric variate.

Note. The Survival function $P\{X > a\} = q^a$, $a = 0, 1, \ldots$ is often useful to recognize geometric distribution. Thus if $X = 1, 2, 3\ldots$, then obviously

$$P(X = a) = P(x > a-1) - P(X > a) = q^{a-1} - q^a = pq^{a-1},\ a = 1, 2, \ldots$$

Example : If $X \sim$ geom (p), show that

(i) $P(X \geq a+b) \mid X \geq a) = P(X \geq b)$, (ii) $P(X = a+b \mid X \geq a) = P(X = b)$. $[a, b = 0, 1, 2, 3, \ldots]$

Solution. (i) For geom (p), the survival function $P(X \geq k) = q^k$, [show it § 14-21]. Hence

$$P(X \geq a+b \mid X \geq a) = \frac{P(X \geq a+b)}{P(X \geq a)} = \frac{q^{a+b}}{q^a} = q^b = P(X \geq b).$$

(ii) $$P(X = a+b \mid X \geq a) = \frac{P(X = a+b)}{P(X \geq a)} = \frac{pq^{a+b}}{q^a} = pq^b = P(X = b).$$

These results refer to memoryless (agelessness) feature of geometric distribution.

14-40. P.G.F., M.G.F. and Moments

$$G(t:X) = p/(1-qt); \; G(t:Y) = pt/(1-qt), \; |qt|<1. \qquad [§ 8\text{-}51]$$

$$M(t:X) = p/(1-qe^t); \; M(t:Y) = pe^t/(1-qe^t), \; |qe^t|<1. \qquad [§ 8\text{-}16]$$

Since $Y = 1 + X$, we could get

$$G(t:Y) = G(t:X+1) = t\,G(t:X) = tp/(1-qt).$$

$$M(t:Y) = M(t:X+1) = e^t M(t:X) = pe^t/(1-qe^t).$$

Cumulants and Moments of Geom (p)

$$K(t:X) = \ln M(t) = \ln[p/(1-qe^t)] = \ln p - \ln(1-qe^t).$$

$$\therefore \quad \sum_{r=1}^{\infty} k_r \frac{t^r}{r!} = \ln p + \sum_{r=1}^{\infty} \frac{q^r e^{rt}}{r}. \qquad \text{[In series expansion]}$$

We differentiate this result n times w.r.t t, and put $t = 0$, to get

$$k_n = \sum_{r=1}^{\infty} q^r r^{n-1} \quad \text{i.e. } k_{n+1} = \sum_{r=1}^{\infty} q^r r^n \qquad \ldots(1)$$

We now differentiate 1(a), the uniformly convergent series, w.r.t. "q" and obtain

$$\frac{dk_n}{dq} = \sum_{r=1}^{\infty} rq^{r-1} r^{n-1} = q^{-1} \sum_{r=1}^{\infty} q^r r^n = q^{-1} k_{n+1} \qquad \text{[by (1)]}$$

$$\therefore \quad k_{n+1} = q\,(dk_n/dq), \; n = 1, 2, \ldots \qquad \text{[Recurrence Formula]} \qquad \ldots(2)$$

From (1) putting $n = 1$, we get

$$k_1 = \sum_{r=1}^{\infty} q^r = \frac{q}{1-q} = \frac{q}{p} \text{ (Mean)}.$$

Various cumulants can now be obtained from (2). Let $D = d/dq$, then

$$k_2 = qD\left(\frac{q}{p}\right) = \frac{q}{p^2}; \quad k_3 = qD\left(\frac{q}{p^2}\right) = \frac{q(1+q)}{p^3}; \quad k_4 = qD\left(\frac{q+q^2}{p^3}\right) = \frac{q(1+4q+q^2)}{p^4}, \text{ etc.}$$

Thus $\mu = k_1 = q/p, \quad \mu_2 = k_2 = q/p^2, \quad \mu_3 = k_3 = q(1+q)/p^3, \quad \mu_4 = k_4 + 3k_2^2,$

$$\beta_1 = \frac{k_3^2}{k_2^3} = \frac{(1+q)^2}{q}, \quad \beta_2 - 3 = \frac{k_4}{k_2^2} = \frac{1+4q+q^2}{q} = \frac{6q+p^2}{q}$$

It may be noted that the point, (β_1, β_2) lies on the St. line $\beta_1 - \beta_2 = 5$.

Observe also : $(k_1/k_2) = p < 1$, $k_2 = k_1 + k_1^2$.

Since Mean $= q/p$ Var $= q/p^2$, hence Var $(X) >$ Mean (X).

Comments. *First Four Moments*

$$K(t) = \ln M(t) = \ln [p/(1-qe^t)] = -\ln \{[1-q(1+t+t^2/2!+\ldots)]/p\}$$

$$= -\ln [1 = \lambda(t+t^2/2!+t^3/3!+\ldots)], \quad (\lambda = q/p)$$

$$= \lambda(t+t^2/2!+\ldots) + \frac{1}{2}\lambda^2(t+t^2/2!+t^3/3!+\ldots)^2 + \frac{1}{3}\lambda^3(t+t^2/2+\ldots)^3 + \frac{1}{4}\lambda^4(t+t^2/2+\ldots)^4\ldots$$

$$k_1 = \lambda = q/p,\ k_2 = \lambda + \lambda^2 = q/p^2, k_3 = \lambda + 3\lambda^2 + 2\lambda^3 = \lambda(1+\lambda)(1+2\lambda) = q(1+q)/p^3, \ \ldots\text{etc.}$$

Factorial m.g.f. and Factorial Moments

$$W(t:X) = \mathrm{E}\{1+t)^X\} = \sum_{x=0}^{\infty} (1+t)^x pq^x = p\sum_{x=0}^{\infty}[q(1+t)]^x = \frac{p}{1-q(1+t)} = \frac{p}{p-qt} = \left[1-\frac{qt}{p}\right]^{-1} \quad \ldots(1)$$

Thus : $$W(t:X) = \sum_{r=0}^{\infty}\left(\frac{qt}{p}\right)^r = \sum_{r=0}^{\infty}\left[\left(\frac{q}{p}\right)^r r!\right]\frac{t^r}{r!}. \qquad \text{[Neg. Binomial Expansion]}$$

$$\therefore \quad \mu'_{(r)} = r!(q/p)^r, \quad r = 0, 1, 2, \ldots$$

We can recover simple moments from Eq. (2).

14-50. *k*-Point Truncated Geometric Distribution

Let $X \sim$ geom (p) and remove probability masses at $x = 0, 1, 2, \ldots k-1$. We need find the distribution of the rest of probability mass.

Observe the Survival function : $P(X \geq a) = q^a$ [§ 14-11]. Now

$$P\{X = x \mid X \geq k\} = \frac{P\{X = x\}}{P(X \geq k)} = \frac{pq^x}{q^k} = pq^{x-k}, x = k, k+1, \ldots$$

If follows that even the truncated distribution is (shifted) geom (p) I $(x \geq k)$.

$$G_X(t) = \mathrm{E}(t^X) = \sum_{x=k}^{\infty} t^x(pq^{x-k}) = pt^k \sum_{y=0}^{\infty}(qt)^y = \frac{pt^k}{1-qt}. \quad [x-k=y]$$

To find the mean and variance of the shelled distribution, we use logarithmic differentiation.

Thus $$G'(t) = \left(\frac{k}{t} + \frac{q}{1-qt}\right)G(t), \quad G''(t) = \left(\frac{k}{t} + \frac{q}{1-qt}\right)G'(t) + \left(\frac{q^2}{(1-qt)^2} = \frac{k}{t^2}\right)G(t)$$

$$\therefore \quad G'(1) = k + (q/p); \quad G''(1) = [k+(q/p)]^2 + [(q^2/p^2) - k]$$

$$\mathrm{E}(X) = G'(1) = k = (q/p), \quad \mathrm{Var}(X) = G''(1) + G'(1) - [G'(1)]^2 = (q/p)^2 + (q/p) = (q/p^2).$$

14-51. Worked-out Problems

***Example* 1.** A man with n keys wants to open his door and tries the keys independently and at random. Find the mean and variance of the number of trials (a) if unsuccessful keys are not eliminated from further selection, and (b) if they are.

Solution. Denote the events : C_j = {success at the jth trial} and let $p_j = P(C_j)$, $q_j = P(\overline{C}_j) = 1 - p_j$.

(a) If the unsuccessful keys are not eliminated, then

$$p_1 = 1/n,\ q_1 = 1 - n^{-1},\ p_2 = 1/n,\ q_2 = 1 - n^{-1}, \ldots,\ p_j = 1/n,\ q_j = 1 - n^{-1}, \ldots$$

The probability of first success is a Pascal geometric variate :

$$f(x) = q^{x-1}\,p,\ x = 1, 2, 3, \ldots \infty.\ \ [p = n^{-1},\ q = 1 - p]$$

$$\therefore \qquad \mathsf{E}(X) = 1/p = n, \mathrm{Var}(X) = q/p^2 = n(n-1).$$

(b) If the unsuccessful keys are eliminated, then denoting the event A_r = {first success at the rth trial}, we have

$$P(C_1) = \frac{1}{n},\ \ P(C_2) = \frac{1}{n-1}, \ldots, P(C_r) = \frac{1}{n-r+1}$$

$$P(A_2) = P(\overline{C}_1\,C_2) = \left(1 - \frac{1}{n}\right)\frac{1}{n-1} = \frac{1}{n};\quad P(A_3) = P(\overline{C}_1\,\overline{C}_2\,C_3) = \left(1 - \frac{1}{n}\right)\left(1 - \frac{1}{n-1}\right)\frac{1}{n-2} = \frac{1}{n},$$

$\ldots P(A_r) = 1/n$. Thus $p_r = P(A_r) = 1/n,\quad \forall\ r = 1, 2, \ldots, n.$

Now if R is the number of trials, then

$$\mathsf{E}(R) = \sum_{r=1}^{n} r\,p_r = \frac{1}{n}\sum_{1}^{n} r = \left(\frac{n+1}{2}\right);\ \ \mathsf{E}(R^2) = \sum_{r=1}^{n} r^2\,p_r = \frac{1}{n}\sum_{1}^{n} r^2 = \frac{(n+1)(2n+1)}{6}.$$

$$\mathrm{Var}\,(R) = \mathsf{E}(R^2) - \mathsf{E}^2(R) = (n^2 - 1)/12.$$

***Example* 2.** *Coupon Collecting Problem.* There are m different types of coupons and each time one gets a coupon it is equally likely to be any of these types. Let X denote the number of coupons one needs to collect so as to possess at least one of each type. Find $\mathsf{E}(X)$ and Var (X).

Let Y denote the number of distinct types of coupons in a collection of n coupons. Find $\mathsf{E}(Y)$ and Var (Y).

Solution. Let X_j denote the number of additional coupons required after j distinct types are collected, until another new type has been obtained, $j = 0, 1, 2, \ldots, m-1$. Then

$$X = X_0 + X_1 + \ldots + X_{m-1}.$$

Note that when j distinct types of coupons have been collected, each new coupon will be a new-type coupon with probability $p_j = (m-j)/m$, so that $X_j \sim$ geom (p_j). Thus

$$\mathsf{E}(X) = \sum_{j=0}^{m-1} \mathsf{E}(X_j) = \sum_{j=0}^{m-1}\left(\frac{m}{m-j}\right) = m\sum_{j=1}^{m}\frac{1}{j}.$$

Further, since each X_j are indep. geom (p_j), we have

$$\text{Var }(X) = \sum_{j=0}^{m-1} \text{Var}(X_j) = \sum_{j=0}^{m-1} \frac{q_j}{p_j^2} = \Sigma\left(\frac{j}{m}\right) \cdot \frac{m^2}{(m-j)^2} = m \sum_{j=0}^{m-1} \frac{j}{(m-j)^2}.$$

Second Part. Let $Y_j = \begin{cases} 1, \text{ if a type } j \text{ coupon is in the collection} \\ 0, \text{otherwise} \end{cases}$

Then $\qquad Y = Y_1 + Y_2 + \ldots + Y_m$

Write $\qquad A_k = \{\text{Collection of } n \text{ coupons contains } \textit{at least} \text{ one type } k \text{ coupon}\}$

$$P_k = P(A_k) = 1 - P(\overline{A}_k) = 1 - \left(\frac{m-1}{m}\right)^n = 1 - \alpha, \quad \left[\alpha - \left(\frac{m-1}{m}\right)^n\right]$$

As $Y_j \sim$ Ber (P_j), we have

$$\text{E}(Y_j) = P_j = 1 - \alpha, \qquad [Q_j = 1 - P_j = \alpha]$$

$$\text{Var }(Y_j) = P_j Q_j = \alpha(1 - \alpha)$$

Thus $\qquad \text{E}(Y) = \sum_{j=1}^{m} \text{E}(Y_j) = m(1-\alpha)$

Observe that $Y_i\, Y_j$ is also Bernoulli (though Y_i, Y_j are dependent) and thus

$$\text{E}(Y_i Y_j) = P\{Y_i Y_j = 1\} = P(A_i A_j) \quad (i \neq j)$$

$$P(A_i A_j) = 1 - P(\overline{A_i A_j}) = 1 - P(\overline{A}_i \cup \overline{A}_j) = 1 - P(\overline{A}_i) - P(\overline{A}_j) + P(\overline{A}_i \,.\, \overline{A}_j)$$

$$= 1 - \alpha - \alpha + \beta.$$

$[\beta \equiv P(\overline{A}_i\, \overline{A}_j) = [(m-2)/m]^n$, collection contains no types i and j coupons]

$\therefore \qquad \text{Cov }(Y_i, Y_j) = \text{E}(Y_i Y_j) - \text{E}(Y_i)\text{E}(Y_j) = (1 - 2\alpha + \beta) - (1-\alpha)^2 = \beta - \alpha^2$

So $\qquad \text{Var }(Y) = \Sigma \text{ Var }(Y_i) + 2\Sigma\Sigma \text{ Cov }(Y_i, Y_j)$

$$= m\alpha(1-\alpha) + 2\binom{m}{2}(\beta - \alpha^2) = m(\alpha - \alpha^2) + m(m-1)(\beta - \alpha^2)$$

$$= m\alpha + m(m-1)\beta - m^2\alpha^2$$

where $\qquad \alpha = [(m-1)/m]^n, \ \beta = [(m-2)/m]^n.$

Problems with Solutions Provided at the End of the Text

1*. If $M_X(t) = (5 - 4e^t)^{-1}$, find $P\{X = 5 \text{ or } 6\}$.

2*. A throws a die repeatedly until he obtains a 6. Independently, B throws a die repeatedly until he obtains a 5 or a 6. If Z is the number of throws required by both A and B together, show that

$$P\{Z = k\} = \frac{5}{18}\left(\frac{5}{6}\right)^{k-2} - \frac{2}{9}\left(\frac{2}{3}\right)^{k-2}, \ k = 2, 3, \ldots$$

3*. Let X and Y be i.i.d. geom (p) and put $Z = \max(X, Y)$. Find $P\{Z = n\}$, $P\{X = k,$ $Z = n\}$, $P\{X = k \mid Z = n\}$, $P\{Z = n \mid X = k\}$, for all integer values $n, k = 0, 1, 2, 3, \ldots$

4*. Show that the generating function of the number of balls that must be distributed to occupy fully the N cells is

$$N!\theta^n(1-\theta)^{-1}(1-2\theta)^{-1}\ldots[1-(N-1)\theta]^{-1}, \quad [\theta = t/N].$$

Exercise 14(a)

1. (a) Two dice are thrown until a seven is obtained. Find the most probable number of throws and also expected number of throws. [Ans. 1 and 6]

   (b) Two dice are thrown until a six is on atleast one of them. Find the chance that for the first time a six appears in the kth throw, $k = 1, 2, 3, \ldots$. [Ans. geom [11/36]]

2. (a) A couple decides to have children until they have a male child. What is the probability distribution of the children they would have ? If the probability of a male child in their community is 1/3, how many children are they excepted to have before the first male child is born ? [Ans. 3]

   (b) Suppose an urn contains 10 balls of which 1 is black. Let Z be the number of the draws with replacement, necessary to observe the black ball. Find the probability distribution of Z and show that $E(Z) = 10$.

3. (a) An urn contains W white and B black balls. Balls are randomly selected, one at a time, until a black one is obtained. Assume that each selected ball is replaced before the next one is drawn. Find the chance that

   (i) exactly n draws are needed, (ii) at least k draws are needed

   (b) In each box of Daffodills-Cosmetics, there is a coupon with a number from 1 to 6. If a person succeeds in getting the series 1 to 6, he gets a gift. How many boxes must he buy, on average, before he gets a gift ? [Ans. 15]

4. (a) A man is given n keys of which only one fits his door. He tries them successively. This may require 1, 2, ..., n trials. Show that each of these n outcomes has probability $1/n$.

   (b) Show that $E(X) = 1.5$, wehre X is a discrete r.v. with $P(X = x) = 2/3^x$, $x = 1, 2, 3, \ldots$

5. Let X be geometric over 0, 1, 2, ... and $f(1) = 3f(2)$. Show that $P(X$ is odd$) = 0.25$

6. Assuming the geometric distribution on positive integers 1, 2, ... find the probability that an integer is (i) even, (ii) odd, (iii) even when it is divisible by 3. [Ans. $q(1+q)^{-1}, (1+q)^{-1}, q^3(1+q^3)^{-1}$]

7. A coin is flipped until the first head occurs. Assume the flips are independent and that the probability of a head occurring each time is p.

   (a) Show that the value of p so that the probability is 0.6 that an odd number of flips is required is 1/3.

   (b) Can you find a value of p so that the probability is 0.5 that an odd number of flips is required ? [No].

8. If X is a geom (p), then $E(X \mid X = n) = E(X) + 1 + n$, for all positive integers n.

9. If $X \sim$ geom (p), then Var $(X \mid X > c) = d$, for $c = -1, 0, 2, \ldots$, and d does not depend on c.

10. Treat gem (p) as a function of p, for fixed x. Show that f is maximum at $p = 1/x$, $x = 2, 3, \ldots$ and that the point of inflection of f is at $p = 2/x$, $x = 3, 4, \ldots$

11. (a) If X and Y are two independent random variables, each representing the number of failures preceding the first success in a sequence of Bernoulli trials with p as probability of success in a single trial and q as a probability of failure, show that $P(X = Y) = p/(1 + q)$

(b) If X and Y are indep. geom (p) variates, show that $P(X = Y) = p/(2 - p)$.

12. Consider a sequence of Bernoulli trials with a constant probability p of success in a single trial. Let X_k denote the number of failures following the $(k - 1)$ and preceding the kth success, and let $S_r = X_1 + X_2 + ... X_r$. Derive the p.d.f. of S_r. Find $E(S_r)$ and Var (S_r).

13. Let X and Y be independent geom (p) variates $W = \min (X, Y)$. Integer $m \geq 0$.

(a) Find the distribution of W. (b) Find $P(W = X)$.

(c) Show that $P(W = m, X - Y = z\} = p^2 (1 - p)^{2m} (1 - p)^{|z|}$ where z is an integer.

(d) Show that min (X, Y) and $X - Y$ are independent.

14. Let X and Y be independent geom (p) variates and $Z = \max (X, Y)$. Find

(a) the distribution of Z. (b) $P(Z = Y)$. (c) the joint distribution of X and Z.

(d) the conditional distn. of X gives $Z = l$. (e) the conditional distn. of Z given $X = k$.

15. Let X and Y be independent geom (p) variates. Let $U = \min (X, Y)$, $V = \max (X, Y)$. Find

(a) $P(U = k, V = k + r)$. (b) $P(U = k)$.

(c) $P(V - U = r)$ (d) Show that U and $V - U$ are independent.

16. Let X and Y be independent geom (p_1) and geom (p_2) variates. Find

(a) $P(X \geq Y)$ (b) $P(X = Y)$ (c) $P\{\min (X, Y) = z\}$ (d) $P(X + Y = z)$.

17. Each of two persons independently throw a coin until he obtains a head. Show that the maximum of throws has the following distribution :

$$p_k = (1/2)^{k-1} - \tfrac{3}{4} (1/4)^{k-1}, \quad k = 1, 2, ...$$

18. A biased coin is tossed indefinitely. Let p $(0 < p < 1)$ be the probability of success. Let Y_1 denote the length of first run and Y_2 the length of the second run. Find the p.m.f.'s of Y_1 and Y_2, and show that $E(Y_1) = (q/p) + (p/q)$, $E(Y_2) = 2$. If Y_n denotes the length of the nth run, $n \geq 1$, what is the p.m.f. of Y_n ? Find $E(Y_n)$.

19. Let X_1, X_2 be i.i.d. random variables having positive mass at 0, 1, 2, ... Also let $U = \max (X_1, X_2)$ and $V = X_1 - X_2$. Then

$$\{P(U = j, V = 0) = P(U = j)\, P(V = 0) \ \forall j\} \Leftrightarrow \{X_1 \text{ and } X_2 \text{ are geometric}\}.$$

20. Let X and Y be mutually independent variables, taking non-negative integer values. Then

$$P(X \leq n) - P(X + Y \leq n) = \lambda P(X + Y = n)$$

holds, for $n = 0, 1, 2, ...$ and same $\lambda > 0$ iff Y is geom $[\lambda/(1 + \lambda)]$.

21. Let $X \sim$ gem (p). Show that $E(X^{-1}) = -\ln [p^{(1/p) - 1}]$

NEGATIVE BINOMIAL DISTRIBUTION

Neg. Bin Thm. $(1-q)^{-k} = \sum_{x=0}^{\infty} \binom{k+x-1}{x} q^x$.

This yields : $\sum_{x=0}^{\infty} \binom{k+x-1}{x} p^k q^x = 1$; $(p+q=1)$, hence the nomenclature

14-60. *Definition of Neg-bin Distribution*

A random variable X governed by the probability law

$$f(x) \equiv P\{X=x\} = \binom{x+k-1}{k-1} p^k q^x,\ (p+q=1),\ x=0,1,2,\ldots,\ f(x)=0,\ \text{otherwise}. \tag{1}$$

Equivalently : $f(x) = P(X=x) = \binom{-k}{x} p^k (-q)^x, 0 \le x < \infty; f(x)=0,$ otherwise

is called a *negative binomial variate*, and the above law is called the *Neg-Bin Distribution*, often indicated by "X is Neg-Bin (k, p)" or X is NB (k, p).

Note. $\binom{x+k-1}{k-1} \equiv \binom{x+k-1}{x} = \dfrac{(x+k-1)(x+k-2)\ldots(k+1)k}{x!}$

$$= \frac{(-1)^x(-k)(-k-1)\ldots(-k-x+2)(-k-x+1)}{x!} = (-1)^x \binom{-k}{x}.$$

This proves the equivalence of two expressions in (1). Obviously,

$$\sum_{x=0}^{\infty} f(x) = \sum_{x=0}^{\infty} \binom{-k}{x} p^k (-q)^x = p^k (1-q)^{-k} = 1.$$

This shows one reason why the distribution is called *negative binomial*.

Bin-Similarity. Let $p = 1/Q$, $q = P/Q$ [or $Q = 1/p$, $P = q/p$]; so that $p + q = 1 \Rightarrow Q - P = 1$. The probability law (1) can then be expressed as a term of *binomial expansion* :

$$P\{X=x\} = \binom{-k}{x} p^k (-q)^x = \binom{-k}{x} Q^{-k-x}(-P)^x,\quad x=0,1,\ldots;\quad f(x)=0,\ \text{otherwise}.$$

In terms of P and Q : $\sum_{x=0}^{\infty} P\{X=x\} = \sum_{x=0}^{\infty} \binom{-k}{x} Q^{-k-x}(-P)^x = (Q-P)^{-k} = 1$.

Comment. $X \sim$ NB $(-k, P)$ if $f(x) = \binom{-k}{x} Q^{-k-x}(-P)^x,\quad x=0,1,2,\ldots$

Physical Conditions for the Occurrence of Negative Binomial Law

The experiments (or trials) with the following properties lead to Neg-Bin distribution :

1. The result of each trial (experiment) can be classified into one of two disjoint categories, say success (S) and failure (F).
2. P (success) $= p$, remains the *same constant* for each trial.

3. The outcomes of all trials are *independent* of each other.
4. The series of trials is performed a *variable* number of times until a *fixed* number k of successes is achieved.

Note. Only the 4th condition differs from the conditions of Binomial distribution

14-61. *Two Forms of Negative Binomial Variate*

Form 1. (Neg-Bin) Let X_k denote the total number of failures preceding the kth success in a sequence of indepe7ndent Bernoulli trials. Then

$$P(X_k = m) = \binom{m+k-1}{k-1} p^k q^m, \quad m = 0, 1, 2, \dots$$

[$X_k = m$ is the number of failures preceding the kth success]

Modelling. Let A = {In the first $m + k - 1$ trials, there are $k - 1$ successes}; B = {mth trial is a success}.

Then, $$P(A) = \binom{m+k-1}{k-1} p^{k-1} q^m \text{ (Binomial law) ;} \quad P(B) = p \text{ (simply).}$$

$$\therefore\ P(AB) = P(A)\,P(B) = \binom{m+k-1}{k-1} p^{k-1} q^m p \Rightarrow f(m) = \binom{m+k-1}{k-1} p^k q^m, \quad m = 0, 1, 2, \dots$$

Note that $f(x) \geq 0, \ \forall \ x$ and $\Sigma f(x) = 1$.

Form 2. (Pascal Neg-Bin) Let Y_k denote the number of trials required upto obtaining the kth success inclusive in a sequence of independent Bernoulli trials. The sample space of Y_k is thus $\{k, k + 1, k + 2, \dots\}$

$$P(Y_k = n) = f(n) = \binom{n-1}{k-1} p^k q^{n-k}, \ n = k, \ k+1, \dots$$

[$Y_k = n$ is the number of trials required to obtain kth success]

Modelling. Let A = {In the first $n - 1$ trials, there are $k - 1$ successes} ; B = {nth trial is a success}.

Then $$P(A) = \binom{n-1}{k-1} p^{k-1} q^{n-k} \quad \text{(Binomial law) ;} \quad P(B) = p \text{ (simply).}$$

$$\therefore \quad P(AB) = P(A)\,P(B) = \binom{n-1}{k-1} p^{k-1} q^{n-k} p \Rightarrow f(n) = \binom{n-1}{k-1} q^{n-k} p^k.$$

Observations. Let $X_k = Y_k - k$, be the number of *failures* upto and including the kth success. Then

$$P\{X_k = r\} = P\{Y_k = r + k\} = \binom{r+k-1}{r} p^k q^r, \text{ for all } r = 0, 1, 2, \dots$$

Remarks and Terminology. The term *modified* negative binomial distribution is often used to describe the variate X_k defined as the total *number of failures before* the kth success in the Bernoulli process. As such, the term neg bin distribution is used to describe the variate Y_k defined as the number of trials necessary to get the kth success ($k \geq 1$). However, we adopt the terminology $X_k \sim$ NB (k, p) and $Y_k \sim$ NB* (k, p) [called

herein Pascal neg-bin distribution] and write X, Y instead of X_k, Y_k. The relation between X_k and Y_k is clearly

$$Y_k = X_k + k, \text{ or simply } Y = X + k.$$

[$X_k + k$ trials are needed to have Y_k failures and k successes]. This relation may be verified by p.g.f. or m.g.f. techniques. Observe that

$$P\{X_k = r\} = P\{Y_k = r + k\} = \binom{r+k-1}{k-1} p^k q^{r-1}, \quad r = 0, 1, \ldots \text{ for each } k = 1, 2, 3, \ldots$$

Comments. N-B distribution also arises when compounding some other probability laws. See Appendix to Chap. 19.

14-70. Neg-Bin Distribution as Convolution of Geom Distribution

Theorem. Let $\{Y_i\}$ be an infinite sequence of (indep.) Bernoulli trials, $P(Y_i = 1) = p$, $P(Y_i = 0) = q = 1 - p$. Let N_r be the number of trials required to obtain the rth success denoted : S_r, $r = 1, 2, \ldots$ The following results hold :

(a) $P(N_1 = k) = pq^{k-1}$, $k = 1, 2, \ldots$ i.e. $N_1 \sim$ geom (p).

(b) Let T_1 = the number of trials required to obtain first success : S_1

T_2 = the number of trials following S_1 upto and including S_2,

....

T_r = the number of trials following S_{r-1} upto and including S_r.

Thus $N_r = T_1 + T_2 + \ldots + T_r$. Show that T_j are independent geom (p) distributed.

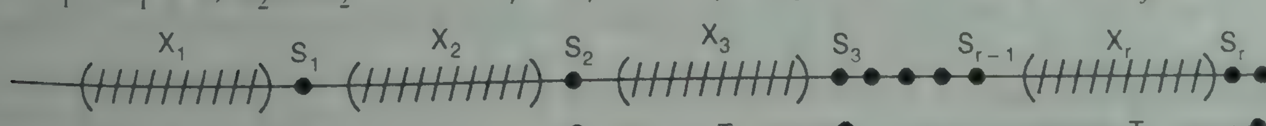
(c) $P(N_r = k) = \binom{k-1}{r-1} p^r q^{k-r} = \binom{-r}{k-1} p^r (-q)^{k-r}, k = r, r+1, \ldots$ i.e. $N_r \sim \text{NB}^*(k, p)$.

Proof. Let X_1 = number of failures preceding the first success S_1

X_2 = number of failures following S_1 and preceding 2nd success S_2, ...

X_r = number of failures between S_{r-1} and S_r.

Thus $T_1 = X_1 + 1$, $T_2 = X_2 + 1$, ..., $T_r = X_r + 1$; $N_r = T_1 + T_2 + \ldots + T_r = (\Sigma X_j) + r$.



(a) $P(N_1 = k) = P\{k - 1 \text{ failures followed by } S_1\} = P(Y_1 = 0, Y_2 = 0, \ldots, Y_{k-1} = 0, Y_k = 1)$

$$= P(Y_1 = 0) \ldots P(Y_{k-1} = 0)\, P(Y_k = 1) = q^{k-1} p, \; k = 1, 2, 3, \ldots \quad \ldots(1)$$

$$G(t : N_1) = \mathrm{E}(t^{N_1}) = \Sigma(q^{k-1} p t^k) = \frac{p}{q} \sum_{k=1}^{\infty} (qt)^k = \frac{pt}{1 - qt}, \quad |qt| < 1 \quad \ldots(2)$$

(b) We prove that T_1 and T_2 are independent. The proof for $T_1, T_2, \ldots, T_r$ is similar but the notation gets involved.

$$P(T_1 = m) = pq^{m-1} \text{ [by (1), since } T_1 = N_1]$$

$$P(T_1 = m, T_2 = n) = P\{Y_1 = \ldots = Y_{m-1} = 0, Y_m = 1; Y_{m+1} = \ldots = Y_{m+n-1} = 0, Y_{m+n} = 1\}$$

$$= (q^{m-1} p)(q^{n-1} p) = p^2 q^{m+n-2}. \qquad \text{[By Independence]}$$

$$P(T_2 = n) = \sum_{m=1}^{\infty} P(T_1 = m, T_2 = n) = p^2 q^{n-2} \sum_{m=1}^{\infty} q^m = pq^{n-1}. \quad \text{[Total Probability Rule]}$$

Thus $P(T_1 = m,\ T_2 = n) = P(T_1 = m)\,P(T_2 = n) \Rightarrow T_1, T_2$ are *i.i.d.* geom (p).

(c) $G(t : N_r) = G(t : T_1 + T_2 + \ldots + T_r) = [G(t : T_1)]^r = [pt/(1-qt)]^r$. [p.g.f. of N - B* (r, p)] ...(3)

$$= p^r t^r (1-qt)^{-r} = p^r t^r \sum_{x=0}^{\infty} (-1)^x \binom{-r}{x} (qt)^x$$

$$\therefore \quad P(N_r = k) = \text{Coeff. of } t^k = \binom{-r}{k-r} p^r (-q)^{k-1} = \binom{k-1}{r-1} p^r q^{k-r}, k = r,\ r+1 \ldots$$

This is Pascal Neg-Bin distribution (r, p) ; rth success on trial k.

Comments. That $T_1, T_2, \ldots, T_r$ are indep. is obvious. In fact, since Bernoulli trials are indep. after achieving S_{j-1}, the process starts over again as a new sequence of Bernoulli trials so that T_j has the same distribution as T_1 which is geom (p). Independence follows since each of r.v.s. T_j is determined by non-overlapping sets of indep. Bernoulli trials, i.e. those trials which occur after S_{j-1} and including the trial at which S_j occurs. We conclude that

$$X = \sum_{j=1}^{r} X_j = \left(\sum_{j=1}^{r} T_j \right) - r = \sum_{j=1}^{r} (T_j - 1). \qquad (T_j - 1) \sim \text{geom } (p) \text{ ; so } X \sim \text{NB } (r, p)]$$

14-71. *Probability Recurrence Formula*

If X is Neg-Bin (k, p) and $f(x)$ is the prob. that there are x failures preceding kth success, then

$$f(x) = \binom{x+k-1}{k-1} p^k q^x \Rightarrow f(x+1) = \binom{x+k}{k-1} p^k q^{x+1}, \quad 0 \le x < \infty.$$

$$\therefore \quad \frac{f(x+1)}{f(x)} = \frac{(x+k)!}{(x+1)!} \frac{x!}{(x+k-1)!} q, \Rightarrow f(x+1) = q \cdot \frac{x+k}{x+1} f(x).$$

This gives the successive probabilities once we know $f(0)$.

Comments. Neg-bin as difference of two sums of Binomial Probabilities

$$p(x) = \binom{k+x-1}{x} p^k q^x, \ x = 0, 1, 2, \ldots \qquad \ldots(1)$$

$P(X > x) = P$ {more than $(x + k)$ trials are required to obtain k successes}

$$= P\{(x+k) \text{ trials produced fewer than } k \text{ successes}\} = \sum_{j=0}^{k-1} \binom{x+k}{j} p^j q^{x+k-j}$$

$$p(x) = P\{X > x-1\} - P(X > x\} = \sum_{j=0}^{k-1} \binom{x+k-1}{j} p^j q^{x-1+k-j} - \sum_{j=0}^{k-1} \binom{x+k}{j} p^j q^{x+k-j}$$

14-72. Worked-out Problems

***Example* 1.** Find the probability that a person tossing 3 coins will get either all heads or all tails for the second time on the fifth toss.

Solution. $p = P\{TTT \uplus HHH\} = (1/8) + (1/8) = 1/4$, thus $p = 1/4$, $q = 3/4$.
Now, the number of trials X on which kth success occurs has the Neg-bin density

$$f(x) = \binom{x-1}{k-1} p^k q^{x-k}, \quad x = k,\ k+1,\ \ldots$$

Here $x = 5$, $k = 2$, $p = 1/4$, hence the required probability P is

$$P = \binom{4}{1}\left(\frac{1}{4}\right)^2 \left(\frac{3}{4}\right)^3 = \frac{3^3}{4^4} = \frac{27}{256}.$$

***Example* 2.** Gamblers A and B, with initial assets a and b, play a game in which, in each play, A beats B with probab. p and loses to B with probab. $q = 1 - p$, $0 < p < 1$. If each play results, in a forfeiture of \$ 1 for the loser and in no change for the winner, find the probability that B will be ruined. [*Attrition Ruin problem by Kaigh, Jan. 1979*]

Solution. Write E_k = {In the first $b + k$ plays, B loses b times}. If every time that A wins is termed a success, then E_k indicates the event that bth success occurs on the $(b + k)$th trial. Thus X is NB distributed. So

$$P(E_x) = \binom{k+b-1}{b-1} p^b q^k, \quad k = 0, 1, 2, \ldots, a-1$$

If A^* denotes that A wins, then $P(A^*) = \Sigma P(E_x)$ and so

$$P(A^*) = \sum_{k=0}^{a-1} \binom{k+b-1}{b-1} p^b q^k.$$

Note. If $a = b = 4$, $p = 0.6$, $q = 0.4$, the $P(A^*) = 0.710$, $P(B^*) = 0.290$.

***Example* 3.** A marksman is required to shoot at a target until he scores 5 bulls' eyes. The probability that he hits the bulls' eye on any trial is 0.3. What is the probability that he requires 8 shots ?

Solution. Here $k = 5$, $p = 0.3$, $x = 3$ (the number of failures preceding the 5th success, as $k + x = 8$). Here by the above relation [§ 14-74]

$$p(x) = \sum_{j=0}^{\infty} \binom{7}{j} (0.3)^j (0.7)^{7-j} - \sum_{j=0}^{\infty} \binom{8}{j} (0.3)^j (0.7)^{8-j}$$

$$= 0.971 - 0.942 = 0.029. \qquad \text{[using binomial tables]}$$

Note. The above solution is only to illustrate above comments. The result follows trivially from (1) :

$$p(x) = {}^7C_3\, (0.3)^5\, (0.7)^3 = 0.029.$$

Problems with Solutions Provided at the End of the Text

1*. If a boy is throwing stones at a target, what is the probability that his 10th throw is his 5th hit, if the probability of hitting the target at any trial is 1/2.

2*. A machine is known to produce 3% defective items. What is the probability that at least 5 items are to be examined in order to get 2 defective items ?

14-80. Generating Functions

$$G_X(t) = \mathrm{E}(t^X) = \sum_{x=0}^{\infty} \binom{k+x-1}{x} p^k q^x . t^x = p^k \sum_{x=0}^{\infty} \binom{k+x-1}{x} (qt)^x$$

$$= p^k (1-qt)^{-k} = [p/(1-qt)]^k$$

$M_X(t)$ has been obtained in § 8-16. This is given by $M_X(t) = [p/(1-qe^t)]^k$.

For the Bin-Similarity, put $p = 1/Q$, $q = P/Q$, $(Q - P = 1)$ $M_X(t) = (Q - Pe^t)^{-k}$.

Example : Show how the moments of Neg-Bin variate can be written down from the corresponding formulas for the binomial variate.

In Bin-Similarity, the comparison is quicker.

Solution.

If Y is bin (n, p) and X is N-B (k, p) then $M(t : Y) = (q + pe^t)^n$, $M(t : X) = (Q - Pe^t)^{-k}$

Thus the correspondence between Y and X is : $q \leftrightarrow Q$, $p \leftrightarrow -P$, $n \leftrightarrow -k$.

The moments relation are thus as under.

Binomial variate. $\mu = np$, $\mu_2 = npq$, $\mu_3 = npq\,(q - p)$, $\mu_4 = npq\,(1 + 3npq - 6pq)$.

Neg-Bin variate. $\mu = kP$, $\mu_2 = kPQ$, $\mu_3 = kPQ\,(Q + P)$, $\mu_4 = kPQ\,[1 + 3kPQ + 6PQ]$.

In terms of p, q. Put $P = q/p$, $Q = 1/p$. Then

Neg-Bin variate : $\mu = \dfrac{kq}{p}$, $\mu_2 = \dfrac{kq}{p^2}$ etc.

14-81. Reproductive Property

Let X_i, $(i = 1, 2, \ldots, n)$ be independent Neg-Bin (k_i, p) variates. Then $S_n = X_1 + X_2 + \ldots + X_n$ is Neg-Bin $(k_1 + k_2 + \ldots + k_n, p)$ variate.

$$M(t : X_1 + X_2 + \ldots + X_n) = M(t : X_1) \,.\, M(t : X_2) \ldots M(t : X_n)$$

$$= \left(\frac{p}{1-qe^t}\right)^{k_1} \cdot \left(\frac{p}{1-qe^t}\right)^{k_2} \ldots, \left(\frac{p}{1-qe^t}\right)^{k_n} = \left(\frac{p}{1-qe^t}\right)^{k}, \quad (k = \Sigma k_j)$$

Thus $\qquad (X_1 + X_2 + \ldots + X_n) \sim \mathrm{NB}\,(k_1 + k_2 + \ldots + k_n, p)$.

14-82. Mass Relations between Binomial and Neg-binomial Distributions

Let $b_n\,(n, p)$ denote that X is bin (n, p) distributed and $\mathrm{NB}_y\,(k, p)$ denote that Y is NB (k, p) distributed. Then the following expressions are true :

(i) $\qquad \mathrm{NB}_y^*(k, p) = \left(\dfrac{k}{y}\right) b_k\,(y, b), \quad y = k, k+1, \ldots$

(ii) $\quad \mathrm{NB}_x(k, p) = \left(\dfrac{k}{k+x}\right) b_k\,(x+k, p),\ x = 0, 1, 2, \ldots$

(iii) $\quad \mathrm{geom}_y\,(p) = (1/y)\, b_1\,(y, p), \quad y = 1, 2, 3, \ldots$

(iv) $\quad \mathrm{geom}_x\,(p) = [1/(1+x)]\, b_1\,(x+1, p),\ x = 0, 1, 2, \ldots$

Proofs are immediate ; e.g. we prove (ii),

$$\frac{k}{k+x}\, b(k : x+k, p) = \frac{k}{k+x}\binom{x+k}{k} p^k q^x = \binom{x+k-1}{k-1} p^k q^x = \mathrm{NB}\,(x ; k, p).$$

Example : Let $f(x) = P\{X = x\}$, $x = 0, 1, 2, \ldots$ be the p.m.f. of a r.v. X. If $(x+1) f(x+1) = (a + bx) f(x)$, $\forall\ x$ and $b \neq -1$, show that

$$\mathrm{E}\{|X - \mu|\} = \frac{2m\, f(m)}{1-b}, \text{ where } \frac{a}{1-b} < m \leq \frac{a}{1-b} + 1,\ (\mu < \infty).$$

Note. Bin (n, p), Pois (λ) and NB (k, p) etc. obey the stated chain relation.

Solution. Here $\sum\limits_0^\infty f(x) = 1, \sum\limits_0^\infty (x+1) f(x+1) = 0\, f(0) + 1 . f(1) + 2f(2) + \ldots = \sum\limits_0^\infty x\, f(x)$.

Randomizing chain relation $\Sigma x f(x) = \Sigma(a + bx) f(x)$ gives

$$\mathrm{E}(X) = a + b\,\mathrm{E}(X) \Rightarrow (1-b)\,\mathrm{E}(X) = a \Rightarrow \mu = a/(1-b). \qquad \ldots(1)$$

We can even find variance, we have

$$\Sigma x(x+1) f(x+1) = a\, \Sigma x f(x) + b\, \Sigma x^2 f(x)$$

i.e. $$\Sigma\,[(x+1)^2 - (x+1)]\, f(x+1) = a\, \Sigma x\, f(x) + b\, \Sigma x^2 f(x)$$

Thus, $\mathrm{E}(X^2) - \mathrm{E}(X) = a\mathrm{E}(X) + b\mathrm{E}(X^2)$

So $$\mathrm{E}(X^2) = (1+a)\,\mu/(1-b) = a(1+a)/(1-b)^2, \qquad \text{[by (1)]}$$

$$\mathrm{Var}\,(X) = \mathrm{E}(X^2) - \mathrm{E}^2(X) = a/(1-b)^2.$$

Mean Absolute Deviation. $M = \mathrm{E}\{|X - \mu|\}$.

$$\mathrm{E}\{|X - \mu|\} = \sum_0^\infty |x - \mu|\, f(x) = \sum_0^{m-1} (\mu - x)\, f(x) + \sum_m^\infty (x - \mu)\, f(x), \qquad [\mu < m < \mu + 1]$$

$$0 = \mathrm{E}\{(x - \mu)\} = \sum_0^{m-1} (x - \mu)\, f(x) + \sum_m^\infty (x - \mu)\, f(x)$$

Adding these equations we get

$$\mathrm{E}\{\,|x - \mu|\,\} = 2 \sum_{x=m}^\infty (x - \mu)\, f(x)$$

or $$M = \frac{2}{1-b} \sum_{x=m}^\infty (x - bx - a)\, f(x) \qquad \left[\text{by (1) } \mu = \frac{a}{1-b}\right]$$

$$\therefore \quad \left(\frac{1-b}{2}\right) M = \sum_{x=m}^\infty [x f(x) - (x+1)\, f(x+1)] \qquad \text{[By Chain relation]}$$

$$= \lim_{n \to \infty} \sum_{x=m}^{m+n} [x\, f(x) - (x+1)\, f(x+1)]$$

$$= \lim_{n \to \infty} \{m\, f(m) - (m+n+1)\, f(m+n+1)\}$$

$$= m\, f(m), [\lim_{n \to \infty} n\, f(n) = 0, \text{ nec. condition for convergence}]$$

Thus $\qquad M = 2m\, f(m)/(1-b).$

Exercise. Extract the value of M for bin (n, p), Pois (λ) and NB (k, p).

14-90. *NB-Distribution as an Extension of Geometric Distribution*

Let Z_1 be the number of trials to the first success and Z_2 be the number of *additional* trials to the second success, Z_3, the number of *additional* trials to the third success, ... and Z_k the number of *additional* trials to the kth success. If Y_k is the number of trials to the kth success, then we must have

$$Y_k = Z_1 + Z_2 + Z_3 + \ldots + Z_k. \qquad \ldots(1)$$

We know that Z_1 is geom (p). So Z_2 must also be geom (p), because the Bernoulli variates involved are independent which guarantee :

$$P(Z_2 = 1) = p,\ P(Z_2 = 2) = pq,\ \ldots,\ P(Z_2 = r) = pq^{r-1}.$$

The same reasoning applies to $Z_3, Z_4, \ldots, Z_k$; i.e. each of Z_j is geom (p) and they are independent. The fact embodied in Eq. (1) then says that the sum of k indep. geom-variates, each with parameter p, is a Neg-bin variate with parameters k and p.

The result also follows by m.g.f. technique. Since Z_i are i.i.d. variates

$$M(t : Y_k) = M(t : Z_1 + Z_2 + \ldots + Z_k) = [M(t : Z_j)]^k = [p/(1 - qe^t)]^k.$$

This shows that $Y_k \sim$ Neg-bin (k, p).

Remarks. Repeated use of Eq. (1) shows that the sum of indep. Neg-bin variates with the same parameter p is again a Neg-bin variate with parameter p. Nicely, for this closure property, m.g.f. method provides the proof instantly.

14-91. *Poisson Distribution as a Limiting Case of Neg-bin Distribution*

If X is Neg-bin (k, p), then as $q \to 0$, $k \to \infty$ such that $kq = \lambda$ (finite), then X tends to be Poisson distributed with parameter λ.

$$M_X(t) = \left(\frac{p}{1 - qe^t}\right)^k = \left(\frac{1 - qe^t}{p}\right)^{-k} = \left(1 - \frac{q(e^t - 1)}{p}\right)^{-k}$$

$$= \left(1 - \frac{\lambda\,(e^t - 1)}{k} \cdot \frac{1}{1-q}\right)^{-k}$$

$$\lim_{k \to \infty} M_X(t) = \lim_{k \to \infty} \left[1 - \frac{\lambda\,(e^t - 1)}{k} \cdot (1-q)^{-1}\right]^{-k} = e^{\lambda(e^t - 1)} \qquad \text{[m.g.f. of Pois } (\lambda)]$$

The unique correspondence between m.g.f. and the p.m.f. now provides that N-B (k, p) tends to be Pois (λ) distributed.

Note. The result can also be established by considering the limiting behaviour of Neg-bin p.m.f.

14-92. *Normal Distribution as a Limiting Case of Neg-bin Distribution*

If X is Neg-bin (n, P), then $M(t : X) = (Q - Pe^t)^{-n}$.

Let $Z = (X - nP) / \sqrt{nPQ}$: where $\mathrm{E}(X) = nP$, Var $(X) = nPQ = \sigma^2$, (Z is standardized). Now

$$M(t : Z) = M[t : (X - nP) / \sigma] = e^{-nPt/\sigma} M[t / \sigma : X]$$

$$= e^{-nPt/\sigma}(Q - Pe^{t/\sigma})^{-n} = (Qe^{Pt/\sigma} - Pe^{Qt/\sigma})^{-n}. \quad \text{...(i)}$$

We now use exponential series to obtain

$$Qe^{Pt/\sigma} - Pe^{Qt/\sigma} = Q\left\{1 + \frac{Pt}{\sigma} + \frac{P^2 t^2}{2\sigma^2} + \frac{P^3 t^3}{6\sigma^2} + \dots\right\} - P\left\{1 + \frac{Qt}{\sigma} + \frac{Q^2 t^2}{2\sigma^2} + \frac{Q^3 t^3}{6\sigma^3} + \dots\right\}$$

$$= 1 - (t^2/2n) - PQ(Q + P)t^3/6\sigma^3 + \dots \quad \text{...(ii)}$$

Substituting from (ii) into (i) we obtain $M(t : Z) = [1 - (t^2 / 2n) + O(1 / n)]^n$.

$$\therefore \quad \lim_{n\to\infty} M(t : Z) = \lim_{n\to\infty}\left[1 - \frac{t^2}{2n} + O\left(\frac{1}{n}\right)\right]^{-n} = e^{t^2/2}. \qquad \text{[Euler limit]}$$

It follows that Z tends to be $N(0, 1)$ distributed when n is large.

14-93. *Mean (Absolute) Deviation*

If X is Neg-bin (k, p), the M.a.D say, M is

$$M = \mathrm{E}(|X - \mu|) = \sum_{x<\mu} f(x).(\mu - x) + \sum_{x>\mu} f(x)(x - \mu), \quad [\mu = kq / p]$$

Also : $$\mathrm{E}(X - \mu) = \sum_{x<\mu} f(x)(x - \mu) + \sum_{x>\mu} f(x)(x - \mu).$$

Adding these we get, using $\lambda = [\mu]$

$$M = 2\sum_{x>\mu} f(x)(x - \mu) = 2\sum_{x=\lambda+1}\binom{k + x - 1}{x} p^k q^x \left(x - \frac{kq}{p}\right)$$

or $$\frac{1}{2}M = \frac{p^{k-1}}{(k-1)!}\sum_{\lambda+1}^{\infty}\frac{(k + x - 1)!}{x!}[xq^x - (x + k)q^{x+1}] \qquad [\because p = 1 - q]$$

$$= \frac{p^{k-1}q}{(k-1)!}\sum_{\lambda+1}^{\infty}\left[\frac{(k + x - 1)!}{(x-1)!}q^{x-1} - \frac{(k + x)!}{x!}q^x\right]$$

$$= \frac{qp^{k-1}}{(k-1)!}\sum_{\lambda+1}^{\infty}(A_{x-1} - A_x) = \frac{qp^{k-1}A_\lambda}{(k-1)!} \qquad \left[A_x = \frac{(k + x)!}{x!}q^x\right]$$

$$\therefore \quad M = 2p^{k-1}q^{\lambda+1}\frac{(k + \lambda)!}{\lambda!(k-1)!} = 2p^{k-1}q^{\lambda+1}(\lambda + 1)\binom{k + \lambda}{\lambda + 1}.$$

Remarks. In allied notation, $p = 1/Q$ and $q = P/Q$, the above result is

$$M = 2P^{\lambda+1}Q^{-\lambda-k}(\lambda + 1)\binom{k + \lambda}{\lambda + 1}, \quad \lambda = [kP].$$

14-94. Worked-out Problems

***Example* 1.** A couple with 2 daughters plans to having children until they have 2 sons. Determine the probability that the couple will have at least 2 more girls before attaining their objective. What shall be expected size of the family then ?

Solution. Let X be the number of additional girls born (i.e. failures) before the couple gets two sons. Then $X \sim$ NB $(2, p)$. So

$$P\{X \geq 2\} = 1 - P\{X < 2\} = 1 - [P(X=0) + P(X=1)]$$

$$= 1 - p^2 - \binom{2}{1} p^2 q = 1 - p^2 - 2qp^2.$$

The expected size of the family, with already 2 girls, is

$$N = 2 + \mathrm{E}(X) + 2 = 4 + 2q/p.$$

Note. If $p = 0.51$ (probab. of a male child), then $P(X \geq 2) = 0.489$, $N = 5.92$.

***Example* 2.** Let X be bin (n, p) and Y be NB (r, p). Show analytically that

$$F_X(r-1) = 1 - F_Y(n-r).$$

Solution. Recall : H-G identity : $\sum_{k=0}^{m} \binom{a}{k}\binom{b}{m-k} = \binom{a+b}{m}$ ...(1)

Negative bin-expansion : $(1-z)^{-N} = \sum_{j=0}^{\infty} \binom{N+j-1}{j} z^j$ ...(2)

$$\therefore \quad H \equiv 1 - F_Y(n-r) = P(Y > n-r) = \sum_{y=n-r+1}^{n} \binom{r+y-1}{r-1} p^r q^y, \qquad [\text{Put } y = (n-r+1)+t]$$

$$= p^r q^{n-r+1} \sum_{t=0}^{\infty} \binom{n+t}{r-1} q^t, \ [\text{Put } p^r q^{n-r+1} = A, \text{ use } a = n, b = t, m = r-1 \text{ in (1)}]$$

$$= A \sum_{t=0}^{\infty} \left\{ \sum_{k=0}^{r-1} \binom{n}{k}\binom{t}{r-1-k} \right\} q^t = A \sum_{k=0}^{r-1} \binom{n}{k} \left\{ \sum_{t=0}^{\infty} \binom{t}{r-1-k} q^t \right\}$$

$$= A \sum_{k=0}^{r-1} \binom{n}{k} \left\{ \sum_{t=r-1-k}^{\infty} \binom{t}{r-1-k} q^t \right\} \qquad \left[\because \binom{t}{j} = 0 \text{ if } j > t\right]$$

$$= p^r q^n \sum_{k=0}^{r-1} \binom{n}{k} q^{-k} \left\{ \sum_{j=0}^{\infty} \binom{r-k+j-1}{j} q^j \right\}, \quad [t = (r-1-k)+j]$$

$$= p^r q^n \sum_{k=0}^{r-1} \binom{n}{k} q^{-k} [(1-q)^{-(r-k)}] = \sum_{k=0}^{r-1} \binom{n}{k} q^{n-k} p^k, \qquad [\text{by (2)}]$$

$$= P\{X \leq r-1\} = F_X(r-1).$$

Remark. Non-analytical proof considers equivalence of event-sets.

Comments. Let $X \sim$ bin (n, p), $Y \sim$ Pas (r, p). Then

(i) $P\{Y \leq n\} = P\{X \geq r\}$, (ii) $P\{Y > n\} = P\{X < r\}$.

***Example* 3.** (*Banach Match-Box Problem*). A pipe-smoking mathematician carries, at all times, two match boxes, one in his left-hand pocket and one in his right-hand pocket. Each time he needs a match he is equality likely to take it from either pocket. Consider the moment when he first discovers that one of his match boxes is empty. Assuming that both match boxes initially contained N matches, find the probability, that there are exactly k matches in the other box $k = 0, 1, 2, \ldots, N$.

Solution. Let A denote the event that the mathematician first discovers that the left-hand match-box is empty (naturally discovered in $N + 1$ trials on L) and there are k matches in the right hand box at that time i.e. $(N - k)$ trials on R. Now event A will occur if the $(N + 1)$th choice of the left-hand match box is made at the $(N + 1) + (N - k)$ trial. Hence from $Y \sim NB^*(r, p)$ with $p = 1/2$, $y = 2N - k + 1$, $k = N + 1$, we get

$$P(A) = \binom{2N-k}{N}\left(\frac{1}{2}\right)^{2N-k+1}$$

If B denotes the event that he first discovers that the right-hand match box is empty and there are k matches in the left-hand box at that time, then by symmetry, $P(B) = P(A)$; hence the probability P of the event of interest is

$$P(A \uplus B) = P(A) + P(B) = 2P(A) = \binom{2N-k}{N}\left(\frac{1}{2}\right)^{2N-k}.$$

Problems with Solutions Provided at the End of the Text

1*. If $X \sim NB(k, p)$, show that the moment recurrence formula is

$$\mu_{r+1} = q\left(\frac{\partial u_r}{\partial q} + \frac{rk}{p^2}\mu_{r-1}\right).$$

2*. If X is Neg-bin (k, Q^{-1}), show that

$$P(X \geq m\} = \frac{1}{B(m,k)}\int_0^P \frac{x^{m-1}\,dx}{(1+x)^{k+m}}, \quad (Q - P = 1).$$

3*. Let $X \sim NB(k, a/(a + t))$. Find the density of **zero-truncated distribution** and express k and a in terms of μ_1', μ_2' and μ_3'.

4*. In a sequence of indep. trials, the probability of a success on each trial is p. By considering the outcome of the first trial, show that $G_r(t)$: the p.g.f. of the number of trials required to achieve the rth success, satisfies $G_r(t) = pt\,G_{r-1}(t) + qt\,G_r(t)$. Hence obtain $G_r(t)$.

Exercise 14(b)

1. State a parametric family of distribution $[\mu = E(X),\ Var(X) = \sigma^2]$ which satisfies :

   (a) $\mu \geq \sigma^2$ (b) $\mu = \sigma^2$ (c) $\mu \leq \sigma^2$ (d) $\mu <, =, > \sigma^2$ (for different values of parameters)

2. To determine who pays for cofee, three people each toss a coin and the odd person pays. If the coins all show heads or all show tails, they are tossed again. What is the probability that a decision is reached in five repetitions or more.

3. Suppose that you believe that you can hit a bull's-eye with a dart 1 time in 10. If this is the case, compute the probability that more than 100 throws are required to obtain exactly 15 bull's eys.
4. An item is produced in large numbers. The machine is known to produce 2% defectives. A quality control insepector is examining the items by taking them at random. What is the probability that at least 4 items are to be examined in order to get 2 defectives ?
5. If $E(X) = 10$ and $\sigma_X = 3$, can X have a N-B distribution ?
6. If X is NB (k, p) and $E(X) = 20$, Var $(X) = 180$, find k, p.
7. If X is Neg-bin variate with parameters k and p, show that,

   (i) $\mu'_{r+1} = (kq/p)\,\mu'_r - q\,(\partial \mu'_r / \partial p)$. (ii) $k_{r+1} = q\,(dk_r / dp)$
8. The number of accidents among 414 machine operators was investigated for three successive months. The following table gives the distribution of the operators according to the number k, of accidents which happened to the same operators. Fit the distribution of the type

   $$P(X = k) = (-1)^r \binom{-n}{k} p^k q^n\ ; \quad k = 0, 1, 2, \ldots n > 0\,; q = 1 - p\,; 0 < p < 1$$

k	0	1	2	3	4	5	6	7	8
Obs. freq.	296	74	26	8	4	4	1	0	1

9. If X is N-B (k, p) and $F(X)$ is the c.d.f. of X, then prove that if (kp/q) is a positive integer, then $F(kp/q) > 1/2$, if $p \le 1/2$.
10. Prove : $P(X \ge m) = \dfrac{1}{B(m, k)} \displaystyle\int_p^1 y^{k-1}(1-y)^{m-1}\,dy$. $[X \sim NB\,(k, p)]$.
11. Let X and Y be independent negative binomial variates with parameters (k_1, P) and (k_2, P) respectively. Show that the conditional distribution of X given $X + Y = n$ is negative hyper-geometric distribution.

$$\binom{x+y}{x} \frac{B(x + k_1, y + k_2)}{B(k_1, k_2)}.$$

You can only be young once. But you can always be immature.

What makes the desert beautiful is that somewhere it hides a well.

More Discrete Distributions 15

15-10. Discrete Uniform Distribution

Definition. Let $A_n = \{a, a+1, a+2, \ldots, b\}$, where $b = a+n-1$, $a \in R$. A r.v. X is said to have a discrete *uniform* (or rectangular) distribution over A_n, written X is $u(A_n)$, if its probability mass function is given by

$$f(x) = 1/n, \quad \forall\ x \in A_n. \quad [n = b-a+1]; \qquad f(x) = 0, \text{ otherwise.}$$

Obviously, $P(X \in B) = \Sigma(1/n)\ \forall\ x \in B,$ $\qquad$ [(points in B)/n]

Remark. This distribution is the basis of equal-likelihood for elementary problems.

15-11. P.G.F., M.G.F., Mean and Variance

$$G(t:X) = \mathrm{E}(t^X) = n^{-1}\,\Sigma t^x = (t^a + t^{a+1} + \ldots + t^{n+a-1})/n$$

$$= t^a(1-t^n)/n(1-t). \qquad [n = b-a+1] \qquad \ldots(1)$$

$$M(t:X) = e^{at}(1-e^{nt})/n(1-e^t). \qquad [t \to e^t \text{ in } (1)] \qquad \ldots(2)$$

$$M(t) = e^{at}(e^{nt}-1)/n(e^t-1)$$

$$= \left(1 + at + \tfrac{1}{2}a^2t^2 + \ldots\right)\left(nt + \tfrac{1}{2}n^2t^2 + \tfrac{1}{6}n^3t^3 + \ldots\right)\left\{n\left(t + \tfrac{1}{2}t^2 + \tfrac{1}{6}t^3 + \ldots\right)\right\}^{-1}$$

$$= \left(1 + at + \tfrac{1}{2}a^2t^2 + \ldots\right)\left(1 + \tfrac{1}{2}nt + \tfrac{1}{6}n^2t^2 + \ldots\right)\left(1 + \tfrac{1}{2}t + \tfrac{1}{6}t^2 + \ldots\right)^{-1}$$

$$= \left[1 + \left(a + \tfrac{1}{2}n\right)t + \left(\tfrac{1}{2}a^2 + \tfrac{1}{2}na + \tfrac{1}{6}n^2\right)t^2 + \ldots\right]\left[1 - \tfrac{1}{2}t - \tfrac{1}{6}t^2 + \left(\tfrac{1}{2}t + \ldots\right)^2 + \ldots\right]$$

$$\mu = \text{Coeff. of } t = a + \tfrac{1}{2}(n-1) = (a+b)/2;\ \mu_2' = \text{Coeff. of } \frac{t^2}{2} = \tfrac{1}{6} - \left(a + \tfrac{1}{2}n\right) + \left(a^2 + na + \tfrac{1}{3}n^2\right)$$

$$\mu_2 = \mu_2' - \mu_1'^2 = \left[a^2 + (n-1)\,a + \tfrac{1}{3}n^2 - \tfrac{1}{2}n + \tfrac{1}{6}\right] - \left[a + \tfrac{1}{2}(n-1)\right]^2$$

$$= \tfrac{1}{3}n^2 - \tfrac{1}{2}n + \tfrac{1}{6} - \tfrac{1}{4}(n-1)^2 = \tfrac{1}{12}(n^2-1) = [(b-a)^2 + 2(b-a)]/12.$$

Direct Evaluation of Mean and Variance

$$\mu = \mathrm{E}(X) = (1/n)\,[a + (a+1) + \ldots + (a+n-1) = a + \tfrac{1}{2}(n-1).$$

$$E(X^2) = (1/n)[\Sigma(a+k-1)^2] = (1/n)\,\Sigma[(a-1)^2 + 2k(a-1) + k^2], \quad (1 \le k \le n).$$

$$= (a-1)^2 + 2(a-1)(\Sigma k)/n + (\Sigma k^2)/n = (a-1)^2 + (a-1)(n+1) + (n+1)(2n+1)/6.$$

$$\text{Var}(X) = E(X^2) - E^2(X) = (a-1)^2 + (a-1)(n+1) + \tfrac{1}{6}(n+1)(2n+1) - [(a-1) + \tfrac{1}{2}(n+1)]^2$$

$$= (1/6)(n+1)(2n+1) - (1/4)(n+1)^2 = (n^2-1)/12.$$

15-12. Worked-out Problems

***Example* 1.** Show that if X, Y be identically distributed with common p.m.f. :

$P(X = k) = 1/N$, for $k = 1, 2, \ldots, N$, $(N > 1)$ [X is uniform on $\{1, 2, \ldots, N\}$]

then $\text{Corr}(X, Y) = 1 - \{6E(X-Y)^2/(N^2-1)\}$.

Solution. Here $E(X) = \sum_{k=1}^{N} k \cdot \dfrac{1}{N} = \dfrac{1}{N}\sum_{k=1}^{N} k = \dfrac{1}{N} \cdot \dfrac{N(N+1)}{2} = \dfrac{N+1}{2} = E(Y)$ [by symmetry]

$$E(X^2) = \Sigma k^2 \cdot (1/N) = (1/N)\,\Sigma k^2 = (N+1)(2N+1)/6$$

$$\text{Var}(X) = E(X^2) - E^2(X) = (N^2-1)/12 = \text{Var}(Y)$$

$$E(X-Y)^2 = E\{(X-\mu_X) - (Y-\mu_Y)\}^2 = E(X_0^2) + E(Y_0^2) - 2E(X_0 Y_0) \quad [X_0 = X - \mu_X, Y_0 = Y - \mu_Y]$$

$$\therefore \quad E(X-Y)^2 = \sigma_X^2 + \sigma_Y^2 - 2\sigma_{XY} = 2(\sigma_X^2 - \sigma_{XY}). \qquad [\sigma_X = \sigma_Y]$$

$$\text{Corr}(X, Y) = \frac{\sigma_{XY}}{\sigma_X \cdot \sigma_Y} = \frac{2\sigma_X^2 - E(X-Y)^2}{2\sigma_X^2} = 1 - \frac{6E(X-Y)^2}{(N^2-1)} = 1 - \frac{6\,E(D^2)}{(N^2-1)}, \quad [D = X - Y].$$

***Example* 2.** A pair of N-sided fair dice is thrown once. If X and Y are the random outcomes on their eyes, find the density function for $Z = X + Y$.

Solution. Here $Z = X + Y$; we need find $P(Z = z)$. If $z = x + y$ and $1 \le x \le N$, $1 \le y \le N$, it follows that $2 \le z \le 2N$. Now

$$P\{X + Y = z\} = \Sigma_x P(X = x, Y = z - x) \qquad \ldots(1)$$

As $\qquad x \ne 0, y \ne 0$ and $y > 0$, so $y = z - x > 0 \Rightarrow x < z \le z - 1$.

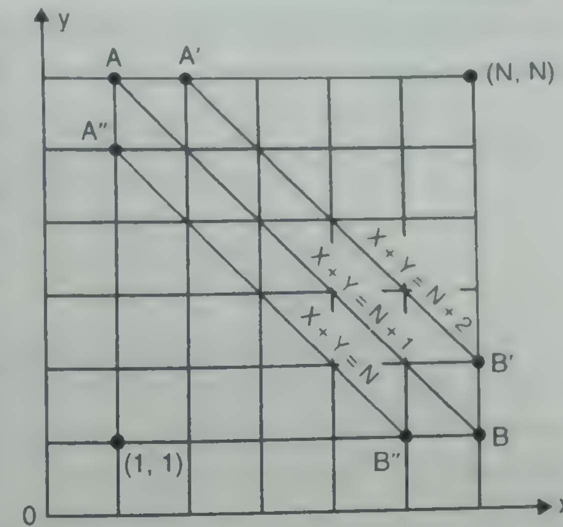
Further, $N = \max y = \max(z - x) = \max z - \min x = \max z - 1$, hence $\max z = 1 + N$, and so $2 \le z \le (1 + N)$. Thus Eq. (1) is valid only over part of the sample space; so

$$P\{X + Y = z\} = \sum_{x=1}^{z-1} P(X = x, Y = z - x) = \sum_{x=1}^{z-1} P(X = x)\,P(Y = z - x) = \sum_{x=1}^{z-1}\left(\frac{1}{N} \cdot \frac{1}{N}\right),$$

where we have used the indep. of X and Y and also equi-probable (discrete uniform) model for calculating probabilities. The constant $(1/N^2)$ occurs $(z - 1)$ times in the above summation, and so

$$P\{X + Y = z\} = (z-1)/N^2, \quad 2 \le z \le (1 + N) \qquad \ldots(2)$$

Now consider the rest of the sample space given by $(N + 2) \le z \le 2N$. We observe that the sample space for $X + Y = Z$ is symmetric about the main diagonal $AB : x + y = N + 1$ (See figure). St. lines $x + y = N + 2$, and $x + y = N$ are reflections of each other in AB and so are $x + y = N + 3$ and $x + y = N - 1$, ... Finally the reflection of (N, N) in AB is the point $(1, 1)$. [Any typical point $z = (x, y) = x + y$ reflects into z', such that $z + z' = 2N + 2$]. The relation of points above and below the diagonal AB is thus $z' = 2(N + 1) - z$. To find p.d.f. for the range $N + 2 \le z \le 2N$, we first observe that the following hold :



$$P\{X + Y = z'\} = (z' - 1\}/N^2, \ 2 \le z' \le (N + 1).$$

Using the above symmetry

$$f(z) = P(Z = z) = P\{X + Y = z\} = \frac{2(N-1) - z - 1}{N^2} = \frac{(2N+1) - z}{N^2}, \quad N + 2 \le z \le 2N.$$

To summarise, the density function of $Z = X + Y$ is given by

$$f(z) = (z - 1)/N^2, \ 2 \le z \le (1 + N); \quad f(z) = [(2N + 1) - z]/N^2, \ N + 2 \le z \le 2N.$$

Note. Probability mass of the diagonal line could be transferred to the range $N + 2 \le z \le 2N$. The density in this case is equivalently given by

$$P\{Z = z\} = \begin{cases} (z-1)/N^2, & 2 \le z \le N \\ (2N + 1 - z)/N^2, & N + 1 \le z \le 2N \end{cases}$$

One can verify that the same results are obtained from the two formulae, etc.

Problems with Solutions Provided at the End of the Text

1*. Let X and Y be i.i.d. uniformly distributed variates over $\{0, 1, 2, ..., N\}$.
(i) Find $P(X \le Y)$, (ii) Find $P(Z = z)$, where $Z = \min(X, Y) = X \wedge Y$.

2*. n observations are drawn independently and at random from a continuous distribution. Find the chance that the next observation is the minimum of all the first n observations.

Exercise 15(a)

1. Let X be uniformly distributed on $\{0, 1, ..., 99\}$. Calculate
   (a) The c.d.f. $F(x)$ (b) $P(2.6 < X < 12.2)$
   (c) $P(25 \le X \le 30)$ (d) $P(X \ge 25)$ (e) $P(8 < X \le 10 \text{ or } 30 < X < 32)$.
2. Let X and Y be i.i.d. $U\{0, 1, 2, ..., N\}$ variates. Find
   (i) $P(X = Y)$ (ii) $P\{\max(X, Y) = z\}$, (iii) $P\{|X - Y| = z\}$.
3. Let X be $U\{0, 1, 2, ..., N\}$. Show that (a) Coeff. of variation $= [(N + 2)/3N]^{1/2}$,
   (b) Coeff. of skewness $= 0$, (c) Coeff. of Excess $= -[1 + 2/N(N + 2)]/5$.
4. If X is uniformly distributed over $\{0, 1, 2, ..., N\}$, prove that

$$\mu_{2r-1} = 0; \mu_{2r} = \frac{1}{N+1}\sum_{j=0}^{N}\left(j - \frac{1}{2}N\right)^{2r}; \ k_{2r-1} = 0, \ r = 1, 2, ...$$

15-20. Hypergeometric Distribution (H-G Distribution)

Definition. A random variable X possessing the probability law :

$$P(X = x) = f(x) = \binom{M}{x}\binom{N-M}{n-x}\Big/\binom{N}{n}, \qquad \max\{0, n-(N-M)\} \le x \le \min\{M, n\}^{*}$$

$$= \binom{Np}{x}\binom{Nq}{n-x}\Big/\binom{N}{n}, \qquad \left[\frac{M}{N} = p, q = 1 - p\right] \qquad \ldots(1)$$

is called a *hyper-geometric* (*H-G*) variate. See Figure below. The distribution (1) is called the hypergeometric distribution. We often write "$X \sim$ H-G (N, M, n)" to indicate that X has the probability law (1). The three constants N, M, n are known as the *parameters of* H-G distribution.

Note. As $x \not> M$ or n, we must have $x \le \min\{M, n\}$. Also $x \ge 0$ and $N - M \ge n - x$, i.e. $x \ge n - N + M$, hence $x \ge \max\{0, M + n - N\}$. We conclude that the possible support of X is

$$\max\{0, M + n - N\} \le x \le \min\{M, n\} \qquad \ldots(2)$$

For x not satisfying (2), $f(x) = 0$ in (1).

Physical Conditions for the Occurrence of H-G Distribution

The trials with the following properties lead to the H-G distribution :

1. The outcome of each trial can be classified into one of two disjoint categories, say success (S) and failure (F).

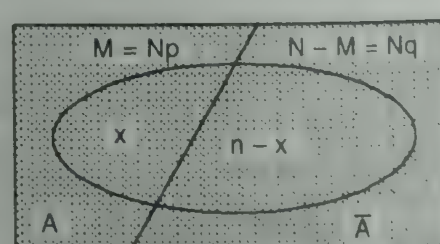
2. Outcomes of successive trials are dependent.

3. The probability of a success changes from trial to trial.

4. The series of trials are performed a *fixed* number of times.

Physical Model for H-G Distribution

From a population of size N (finite number) M elements are of type A and $N - M$ are those of type A'. Given that n elements are withdrawn at random from this population *without replacements*, show that the probability that exactly x of the n elements are of type A is



HYPERGEOMETRIC MODEL

$$f(x) = \binom{M}{x}\binom{N-M}{n-x}\Big/\binom{N}{n}.$$

Proof. From the figure, we observe that there are ${}^{M}C_x$ ways of choosing x elements from M elements composed of the sub-population A, followed by ${}^{(N-M)}C_{(n-x)}$ independent ways of choosing the remaining $(n - x)$ elements from the sub-population A'. By the Principle of Sequential Counting, the total number of required selections is

$\binom{M}{x}\binom{N-M}{n-x}$. The exhaustive number of possible withdraws is $\binom{N}{n}$. Thus, the probability mass is,

---

* We can take $x = 0, 1, 2, \ldots, n$, because $P\{x = n\} = 0$, unless $x \in \min\{M, n\}$, because $\binom{r}{k} = 0$ if $k > r$ or if $k < 0$

$$f(x) = \binom{M}{x}\binom{N-M}{n-x}\Big/\binom{N}{n}, \quad x = 0, 1, 2, \ldots$$

Remarks. (i) $f(x) \geq 0$. (ii) $\Sigma f(x) = 1, \ \forall \ x \in [\max\{0, M+n-N\}, \min\{M, n\}]$
The (ii) follows instants from the following H-G identity

$$\sum_{x=0}^{n}\binom{M}{x}\binom{N-M}{n-x} = \binom{N}{n}$$

15-21. *Structural Similarity between Bin and H-G Probabilities*

$$f(x) = \binom{Np}{x}\binom{Nq}{n-x}\Big/\binom{N}{n}, \qquad x = 0, 1, 2, \ldots, n$$

Use : $\binom{N}{k} = \dfrac{N!}{k!(N-k)!} = \dfrac{N^{(k)}}{k!}$. Then

$$f(x) = \frac{(Np)^{(x)}}{x!}\frac{(Nq)^{(n-x)}}{(n-x)!}\cdot\frac{n!}{N^{(n)}}. = \binom{n}{x}\frac{(Np)^{(x)}(Nq)^{(n-x)}}{N^{(n)}}. \quad \text{[Factorial exponents]} \quad \ldots(1)$$

$$b(x) = \binom{n}{x}p^x q^{n-x} = \binom{n}{x}\frac{(Nq)^{n-x}.(Nq)^x}{(N)^n}, \quad x = 0, 1, 2, \ldots, n \quad \ldots(2)$$

Thus, ordinary exponents, when changed to factorial exponents, bin (n, p) converts to H-G (N, M, n), $(p = M/N)$ and conversely.

15-22. *Probability Mass Recurrence Formula*

If X is Hyp-geom (N, M, n), we have

$$f(x) = \binom{M}{x}\binom{N-M}{n-x}\Big/\binom{N}{n}; \quad f(x+1) = \binom{M}{x+1}\binom{N-M}{n-x-1}\Big/\binom{N}{n}$$

Dividing : $$\frac{f(x+1)}{f(x)} = \frac{\binom{M}{x+1}\binom{N-M}{n-x-1}}{\binom{M}{x}\binom{N-M}{n-x}} = \frac{(M-x)}{(x+1)}\frac{(n-x)}{(N-M-n+x+1)}$$

Thus $$f(x+1) = \frac{(n-x)}{(x+1)}\cdot\frac{M-x}{(N-M-n+x+1)}f(x). \quad \ldots(1)$$

Note. If $N \to \infty$. $(M/N) \to p$, this formula gives

$$f(x+1) = \frac{n-x}{x+1}\cdot\frac{p}{q}f(x) \quad \text{[Bin-density recurrence formula]} \quad \ldots(2)$$

From (1) and (2), it follows that Hyp-geom $(N, M, n) \to$ bin (n, p) under above conditions.

15-23. *Modal Value*

For the positive integers N, M, n, let $\theta = (n + 1)(M + 1)/(N + 2)$ ...(1)

and let ζ be the unique *integer* satisfying $\theta - 1 \le \zeta \le \theta$, i.e. $\zeta \le \theta \le \zeta + 1$.

(a) If $\zeta = \theta$ (i.e. θ itself is an integer), then the H-G densities satisfy

$$p_0 < p_1 < \ldots < p_{\zeta-1} \quad = \quad p_\zeta, \text{ and } \quad p_\zeta > p_{\zeta-1} > \ldots > p_m. \qquad [m = \min(M, n)]$$

(b) If $\zeta \ne \theta$ (i.e. θ is not an integer), then

$$p_0 < p_1 < \ldots < p_\zeta \text{ and } \quad p_\zeta > p_{\zeta-1} > \ldots > p_m.$$

In either case $\qquad |(\zeta / n) - (M / n)| < 1 / n.$ ...(2)

Proof. By Mass Recurrence Formula [§ 15-23 (1)], $p_k = f(k)$

$$\frac{f(k+1)}{f(k)} = \frac{(M-k)(n-k)}{(k+1)(N-M-n+k+1)}. \qquad \text{...(3)}$$

Hence, $[f(k + 1)/f(k)] > 1$ yields

$$(M-k)(n-k) > (k+1)[N-M+1-(n-k)] \Rightarrow (N+2)k < (n+1)(M+1) - (N+2)$$

i.e. $$k < [(n + 1)(M + 1)/(N + 2)] - 1 = \theta - 1.$$

Thus, $p_k < p_{k+1}$, for all $k < \theta - 1$. By considering $p_{k+1} < p_k$ in (3), we get that $p_k > p_{k+1}$, $\forall\, k > \theta - 1$, of course, with $k < \min(M, n)$. We have thus proved that if θ is an integer

$$p_0 < p_1 < p_2 < \ldots p_{\theta-1}, \text{ and } \quad p_\theta > p_{\theta+1} > \ldots > p_m.$$

If θ is not an integer, then with the unique integer ζ as defined,

$$p_0 < p_1 < \ldots < p_\zeta, \text{ and } \quad p_\zeta > p_{\zeta+1} > \ldots > p_m$$

because if θ is not an integer, the relations $k < \theta - 1$ and $k > \theta - 1$ are equivalent to $k \le \zeta - 1$ and $k \ge \zeta$, respectively. Consequently, to prove the result we need to show that $p_{\theta-1} = p_\theta$, whenever θ is an integer. For this, put $k = \theta - 1$ in (3) to get

$$\frac{p_\theta}{p_{\theta-1}} = \frac{(M-\theta+1)(n-\theta+1)}{\theta(N-M-n+\theta)} \qquad \text{...(4)}$$

Putting for θ from (1) in various terms of (4), we observe that

$$M+1-\theta = (M+1)(N-n+1)/(N+2); \qquad (n+1)-\theta = (n+1)(N-M+1)/(N+2)$$

$$N-M-n+\theta = (N-n+1) - [(M+1)-\theta] = (N-n+1)(N-M+1)/(N+2).$$

Substituting for θ and the above three evaluations in (4), we find that $(p_\theta / p_{\theta-1}) = 1$, i.e. $p_\theta = p_{\theta-1}$, as was claimed.

To estimate the quantity : $(\zeta / n) - (M / N)$, we observe by minor algebra

$$\zeta \le \theta \le \zeta + 1 \Rightarrow (\zeta / n) - (M / N) \le (\theta / n) - (M / N) \le (\zeta / n) + (1 / n) - (M / N)$$

We rearrange this chain as

$$(\theta / n) - (M / N) - (1 / n) < (\zeta / n) - (M / N) \le (\theta / n) - (M / N) \qquad \ldots(5)$$

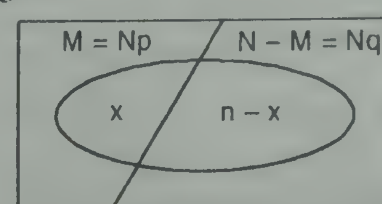
$$\text{(i)} \qquad\qquad \text{(ii)} \qquad\qquad \text{(iii)}$$

$$\text{(iii)} = \frac{N(M+n+1) - 2Mn}{nN(N+2)} = \frac{n(N-M) + M(N-n) + N}{n(N^2+2N)} \qquad \ldots(6)$$

$$\le \frac{nN + N(N-n) + N}{n(N^2+2N)} = \frac{N^2+N}{n(N^2+2N)} < \frac{1}{n}.$$

Obviously, (iii) is positive [see last fraction in (6)] and less than $(1/n)$. Also, (i) is smaller by $(1/n)$ than (iii). Hence, the above estimate of (i) implies that $|\,(i)\,| < (1/n)$. And this proves our assertion (2).

15-24. *Simple Factorial Moments. Mean and Variance*



$$f(x) = \binom{Np}{x}\binom{Nq}{n-x}\Big/\binom{N}{n}, \quad x = 0, 1, 2, \ldots, n$$

$$\therefore \quad \mathrm{E}\{X^{(r)}\} = \sum_x f(x) \, . \, x^{(r)}$$

$$= \sum_x x^{(r)} \, . \, \frac{(Np)^{(r)}}{x^{(r)}} \binom{Np-r}{x-r}\binom{Nq}{n-x}\Big/\binom{N}{n}$$

$$= (Np)^{(r)} \sum_{x=r}^{n} \binom{Np-r}{x-r}\binom{Nq}{n-x}\Big/\binom{N}{n}$$

$$= (Np)^{(r)} \left\{\sum_{j=0}^{n-r} \binom{Np-r}{j}\binom{Nq}{n-r-j}\right\}\Big/\binom{N}{n}, \qquad [j = x - r]$$

We apply H-G identity and also replace $\binom{N}{n}$ by $\dfrac{N^{(r)}}{n^{(r)}}\binom{N-r}{n-r}$. Thus

$$\mathrm{E}\{X^{(r)}\} = \left\{\frac{(Np)^{(r)}\, n^{(r)}}{N^{(r)}} \cdot \frac{1}{\binom{N-r}{n-r}}\right\}\left\{\binom{Np+Nq-r}{n-r}\right\}$$

$$\therefore \quad \mu'_{(r)} \equiv (Np)^{(r)} \, . \, n^{(r)} / N^{(r)} \qquad \ldots(1)$$

$$\mathrm{E}(X) = \mu'_{(1)} = (Np)n / N = np$$

$$\mathrm{E}[X(X-1)] = \frac{(Np)(Np-1) \, . \, n(n-1)}{N(N-1)} \Rightarrow \mathrm{E}(X^2) = np + \left\{\frac{np\,(Np-1)\,(n-1)}{(N-1)}\right\}$$

$$\text{Var}(X) = \mathsf{E}(X^2) - \mathsf{E}^2(X) = np\left\{1 + \frac{(Np-1)(n-1)}{N-1} - np\right\}$$

$$= \frac{np}{N-1}(Nq - nq) = \frac{npq}{N-1}(N-n) = (npq)\left\{1 - \frac{(n-1)}{(N-1)}\right\}.$$

Cor. Factorial moments of bin (n, p) from HG(N, M, n)

$$\mu'_{(r)} = n^{(r)} \cdot \frac{Np(Np-1)\dots(Np-r+1)}{N(N-1)\dots(N-r+1)}$$

$$= n^{(r)}\left\{p\left(p-\frac{1}{N}\right)\left(p-\frac{1}{N}\right)\dots\left(p-\frac{r-1}{N}\right)\right\} \Big/ \left\{\left(1-\frac{1}{N}\right)\left(1-\frac{2}{N}\right)\dots\left(1-\frac{r-1}{N}\right)\right\}$$

Let $N \to \infty$, then $\lim\limits_{n\to\infty} \mu'_{(r)} = n^{(r)} p^r$

Note. $$N^{(k)} = \frac{N!}{(N-k!)} = \frac{(\text{Base})!}{(\text{Base - Power})!}.$$

Illustration : Find the S.D. of the dist $f(x) = {}^4C_x\ {}^6C_{6-x} / {}^{10}C_4, x = 0, 1, \dots, 4.$ Here

$$X \sim \text{Hyp-geom } (10, 4\ ;\ 6)\ ;\ \text{so } \sigma^2 = 64/100.$$

15-25. *Relation of H-G Distribution to Hypergeometric Series & P.G.F.*

The definitions and some relations between increasing and decreasing factorial powers are :

Def. **:** $n^{(k)} = n(n-1)\dots(n-k+1);$ $\qquad n^{[k]} = n(n+1)(n+2)\dots(n+k-1).$

Reln. **:** $n^{[k]} = (-1)^k(-n)^{(k)}, \quad n^{[k]} = (n+k-1)^{(k)}, \quad N^{(n-x)} = N^{(n)} / (N-n+x)^{(x)}.$ ...(1)

The p.m.f. of H-G distribution can be converted into decreasing factorial powers :

$$f(x) = n^{(x)} . (Np)^{(x)} . (Nq)^{(n-x)} / x!\, N^{(n)}. \qquad [\S\ 15\text{-}21(1)] \qquad ...(2)$$

The Hypergeometric series is

$$F\left(\begin{matrix}\alpha, \beta\\ \gamma\end{matrix}\middle| t\right) = 1 + \frac{\alpha\beta}{\gamma}t + \frac{\alpha(\alpha+1)\beta(\beta+1)}{\gamma(\gamma+1)}\frac{t^2}{2!} + \dots = 1 + \sum_{x=1}^{\infty}\frac{\alpha^{[x]}\beta^{[x]}}{\gamma^{[x]}}\cdot\frac{t^x}{x!} \qquad ...(3)$$

Put $\alpha = -n, \beta = -Np, \gamma = Nq - n + 1;$ then using (1) the Eq. (3) reduces to

$$F\left(\begin{matrix}\alpha, \beta\\ \gamma\end{matrix}\middle| t\right) = 1 + \sum_{x=1}^{\infty}\frac{(-n)^{[x]}.(-Np)^{[x]}}{(Nq-n+1)^{[x]}}\frac{t^x}{x!} = 1 + \sum_{x=1}^{\infty}\frac{n^{(x)}(Np)^{(x)}}{(Nq-n+x)^{(x)}}\frac{t^x}{x!} \qquad \text{[by 1(a) \& 1(b)]}$$

$$\therefore F\left(\begin{matrix}-n,\ Np\\ Nq-n+1\end{matrix}\middle| t\right) = 1 + \sum_{x=1}^{n}\frac{n^{(x)}.(Np)^{(x)}.(Nq)^{(n-x)}}{(Nq)^{(n)}}\frac{t^x}{x!} = \sum_{x=0}^{\infty}\frac{N^{(n)}}{(Nq)^{(n)}}f(x)t^x. \qquad \text{[by 1(c) \& (2)]}$$

$$\therefore \quad G_X(t) = \sum_{x=0}^{\infty} f(x) . t^x = \frac{(Nq)^{(n)}}{N^{(n)}} \; F\left(\begin{matrix} -n, -Np \\ Nq - n + 1 \end{matrix} \middle| \, t \right). \quad \text{...(4)}$$

Thus, $f(x)$ is the coefficient of t^x in a constant multiple of H-G Series.
Incidently, (4) is the p.g.f. of H-G (Np, Nq, n) distribution. Obviously

$$M(t : X) = \frac{(N-M)!(N-n)!}{N!} F\left(\begin{matrix} -n, -M \\ N - M - n + 1 \end{matrix} \middle| \, e^t \right). \qquad [\text{Change } t \text{ to } e^t \text{ in (4)}] \quad \text{...(5)}$$

15-26. *Approximation to Binomial Distribution*

When $N \to \infty$ and $(M/N) = p$, H-G density tends to binomial density.

Proof. Here $$f(x) = \binom{Np}{x}\binom{Nq}{n-x} \bigg/ \binom{N}{n}. \quad \text{...(1)}$$

$$\binom{Np}{x} = \frac{Np\,(Np-1)\ldots(Np-x+1)}{x!} = \frac{pN^x}{x!}\left(p - \frac{1}{N}\right)\left(p - \frac{2}{N}\right)\ldots\left(p - \frac{x-1}{N}\right) \quad \text{...(2)}$$

$$\binom{Nq}{n-x} = \frac{q(N)^{n-x}}{(n-x)!}\left(q - \frac{1}{N}\right)\left(q - \frac{2}{N}\right)\ldots\left(q - \frac{n-x-1}{N}\right) \quad \text{...(3)}$$

$$\binom{N}{n} = \frac{N(N-1)\ldots(N-n+1)}{n!} = \frac{(N)^n}{n!}\left(1 - \frac{1}{N}\right)\left(1 - \frac{2}{N}\right)\ldots\left(1 - \frac{n-1}{N}\right) \quad \text{...(4)}$$

Substituting from (2), (3) and (4) into (1) we get

$$f(x) = \frac{n!\,pq}{x!(n-x)!} \; \frac{\left(p - \frac{1}{N}\right)\left(p - \frac{2}{N}\right)\ldots\left(p - \frac{x-1}{N}\right).\left(q - \frac{1}{N}\right)\left(q - \frac{2}{N}\right)\ldots\left(q - \frac{n-x-1}{N}\right)}{(1-N^{-1})(1-2N^{-1})\ldots[1-(n-1)N^{-1}]} \quad \text{...(5)}$$

$$\therefore \quad \lim_{n \to \infty} f(x) = \binom{n}{x} q^{n-x} p^x, \; 0 \le x \le n.$$

Note. In sampling with replacement, the probability of successive successes decreases.

15-27. *Worked-out Problems*

***Example* 1.** Let X_j, $1 \le j \le n$ be i.i.d. Indicator (i.e. Bernoulli) variates, $P(X_j = 1) = p$, $P(X_j = 0) = q = 1 - p$. If $S_t = X_1 + X_2 + \ldots + X_t$, find $P\{S_k = r \mid S_n = m\}$.

Solution. Here $G(t : X_1) = \mathrm{E}(t^{X_1}) = q + pt$; $G(t : S_k) = (q + pt)^k$; so $S_k \sim$ bin (k, p). Now, $S_n = m$ gives

$$(X_1 + \ldots + X_k) + (X_{k+1} + \ldots + X_n) = S_k + S_{n-k} = m.$$

There are r successes in k trials [$S_k \sim$ bin (k, p)] and $m - r$ successes in $n - k$ trials, [$S_{n-k} \sim$ bin $(n - k, p)$], hence by independence of S_k and S_{n-k} binomial variates

$$P\{S_k = r, S_{n-k} = m-r\} = P(S_k = r)\, P(S_{n-k} = m-r)$$

$$= \binom{k}{r} q^{k-r} p^r \cdot \binom{n-k}{m-r} p^{m-r} q^{n-k-m+r} = \binom{k}{r}\binom{n-k}{m-r} p^m q^{n-m}. \qquad \ldots(1)$$

$$\therefore \quad P\{S_k = r \mid S_n = m\} = P\{S_k = r, S_{n-k} = m-r\} / P(S_n = m)$$

$$= \binom{k}{r}\binom{n-k}{m-r} p^m q^{n-m} \Big/ \binom{n}{m} q^{n-m} p^m = \binom{k}{r}\binom{n-k}{m-r} \Big/ \binom{n}{m}, \; 1 \le k \le n; 0 \le r \le m \le n.$$

We observe that the conditional probability is independent of p and obeys H-G probability law.

Remarks. The following example is a repetition of above in a popular brand of exercises.

***Example* 2.** If X and Y are independent binomial variates with parameters n_1, p and n_2, p; find $P\{X = r \mid X + Y = k\}$.

Solution. Recall that $X + Y \sim$ bin $(n_1 + n_2, p)$. Now using Indep. (X, Y)

$$P\{X = r \mid X + Y = k\} = \frac{P(X = r, X+Y = k)}{P(X+Y = k)} = \frac{P(X = r, Y = k-r)}{P(X+Y=k)} = \frac{P(X=r)\, P(Y = k-r)}{P(X+Y=k)}$$

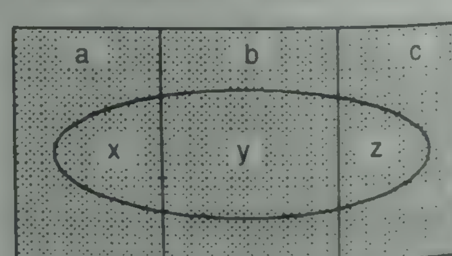
$$= \frac{{}^{n_1}C_r\, p^r q^{n_1 - r} \cdot {}^{n_2}C_{k-r}\, p^{k-r} q^{n_2 - k + r}}{{}^{(n_1+n_2)}C_k\, p^k q^{n_1 + n_2 - k}} = \binom{n_1}{r}\binom{n_2}{k-r} \Big/ \binom{n_1 + n_2}{k} : \text{indep. of } p.$$

Remarks. The conditional distribution of X given $X + Y = k$ is hyper-geometric. The problem can be solved, without calculations, by additive arguments.

***Example* 3.** Find the Coeff. of correlation in a bivariate hypergeometric distribution.

Solution. $P(X = x, Y = y, Z = z) = \binom{a}{x}\binom{b}{y}\binom{c}{z} \Big/ \binom{N}{n}$. $\quad \begin{cases} N = a + b + c \\ n = x + y + z \end{cases}$

There are only two independent variables, X and Y, since $Z = n - X - Y$. Now



BIV. H-G MODEL

Recall : $\quad \binom{M}{K} = \frac{M^{(r)}}{K^{(r)}} \binom{M-r}{K-r};$

$$\mathrm{E}[X^{(r)} Y^{(s)} Z^{(t)}] = \binom{N}{n}^{-1} \sum_{x, y, z} x^{(r)} \binom{a}{x} y^{(s)} \binom{b}{y} z^{(t)} \binom{c}{z}$$

$$= \binom{N}{N}^{-1} a^{(r)} b^{(s)} c^{(t)} \Sigma \binom{a-r}{x-r}\binom{b-s}{y-s}\binom{c-t}{z-t} \qquad \ldots(1)$$

Now, by repeated use of Hypergeometric Identity :

$$\sum_{x=r}^{a} \binom{a-r}{x-r} \sum_{y=s}^{b} \binom{b-s}{y-s}\binom{N-a-b-t}{n-x-y-t} = \sum_{x=r}^{a} \binom{a-r}{x-r}\binom{N-a-s-t}{n-x-s-t} = \binom{N-r-s-t}{n-r-s-t}.$$

Consequently, if $r + s + t = m$, then

$$\mathrm{E}[X^{(r)} Y^{(s)} Z^{(t)}] = \left[\binom{N-m}{n-m} \Big/ \binom{N}{n}\right] a^{(r)} b^{(s)} c^{(t)} = \frac{n^{(m)}}{N^{(m)}} \cdot a^{(r)} b^{(s)} c^{(t)}.$$

Thus, $\mathsf{E}(X) = \dfrac{na}{N} = np_1 \,;\mathsf{E}(Y) = \dfrac{nb}{N} = np_2 \,;\mathsf{E}(XY) = \dfrac{n^{(2)}ab}{N^{(2)}}.$ $\left[p_1 = \dfrac{a}{N}, p_2 = \dfrac{b}{N}\right]$

$$\mathsf{E}[X^{(2)}] = \frac{n^{(2)}\,a^{(2)}}{N^{(2)}} \Rightarrow \mathsf{E}(X^2) = \frac{n(n-1)\,a(a-1)}{N(N-1)} + \frac{na}{N}$$

So $\text{Var}\,(X) = \mathsf{E}(X^2) - \mathsf{E}^2(X) = \dfrac{na}{N}\,\dfrac{(N-n)\,(N-a)}{N(N-1)} = \dfrac{np_1\,q_1(N-n)}{N-1}.$

Similarly, $\text{Var}\,(Y) = np_2\,q_2\,(N-n)/(N-1).$

$$\sigma_{XY} = \mathsf{E}(XY) - \mathsf{E}(X)\,\mathsf{E}(Y) = \frac{n(n-1)\,ab}{N(N-1)} - \frac{n^2ab}{N^2} = -np_1\,p_2\left(\frac{N-n}{N-1}\right). \quad \ldots(2)$$

Thus, $\rho\,(X,Y) = \dfrac{\sigma_{XY}}{\sigma_X\,.\,\sigma_Y} = -\dfrac{p_1\,p_2}{\sqrt{p_1\,q_1\,p_2\,q_2}} = -\sqrt{\dfrac{p_1\,p_2}{q_1\,q_2}}.$

Note. Compare (2) with §-15-35 (i) for Trinomial Distribution.

***Example* 4.** An urn contains w white balls and b black balls. Balls are drawn one at a time from the urn, without replacement. Find the distribution of the number X of draws needed to obtain the kth black ball. Find also $\mathsf{E}(X^{[m]})$. (*Negative hyper-geom distribution*).

Solution. We observe the trivial fact.

$P\{k$th black ball on the xth darw$\} = P\{(k-1)$ black balls in the first $(x-1)$ draws followed by 1, black ball from $[b-(x-1)]$ black and $[w-(x-k)]$ white balls$\}$

$$\therefore \quad f(x) = P(X=x) = \frac{\dbinom{b}{k-1}\dbinom{w}{x-k}}{\dbinom{b+w}{x-1}} \cdot \frac{b-(k-1)}{(b+w-x+1)}$$

Expressing the bin-coeffts in terms of factorials and regrouping the factorials to form bin-coeffts, we get

$$f(x) = \binom{x-1}{k-1}\binom{b+w-x}{b-k}\Big/\binom{b+w}{b} = \binom{x-1}{k-1}\binom{N-x}{b-k}\Big/\binom{N}{b},\ x \approx k,\, k+1,\, \ldots\, (b+w) \quad \ldots(1)$$

By Normality of p.m.f., viz. $\sum f(x) = 1$, This gives

$$\sum_{x=k}^{N}\binom{x-1}{k-1}\binom{N-x}{b-k} = \binom{N}{b}, \qquad N = (b+w) \qquad \ldots(2)$$

The *advancing* mth factorial moment is

$$\mathsf{E}\{X^{[m]}\} = \sum_{x=k}^{N} x^{[m]}\binom{x-1}{k-1}\binom{N-x}{b-k}\Big/\binom{N}{b}, \qquad \left\{x^{[m]} = x(x+1)\ldots(x+m-1)\right\}$$

$$= (k)^{[m]}\sum_{x<k}^{N}\binom{x+m-1}{k+m-1}\binom{N-x}{b-k}\Big/\binom{N}{b}, \qquad \left\{x^{[m]}\binom{x-1}{k-1} = k^{[m]}\binom{x-1+m}{k-1+m}\right\}$$

$$= (k)^{[m]} \sum_{z=k+m}^{N+m} \binom{z-1}{k+m-1}\binom{N+m-z}{(m+b)-(k+m)} \Big/ \binom{N}{b}, \qquad \{z = x+m\}$$

$$= (k)^{[m]} \binom{N+m}{b+m} \Big/ \binom{N}{b}, \qquad \text{[Change } N \text{ to } N+m,\ b \text{ to } b+m \text{ in (2)]}$$

$$= (k)^{[m]} \frac{(N+m)!}{(b+m)!}\,\frac{b!}{N!} \qquad \text{...(3)}$$

This gives $\quad \mathrm{E}(X) = \dfrac{k(N+1)}{(b+1)}, \ \mathrm{E}(X^2) = \dfrac{k(k+1)(N+1)(N+2)}{(b+2)(b+1)} - \dfrac{k(N+1)}{(b+1)}.$

$$\mathrm{Var}\,(X) = \frac{kw(N+1)(b+1-k)}{(b+1)^2\,(b+2)}.$$

Problems with Solutions Provided at the End of the Text

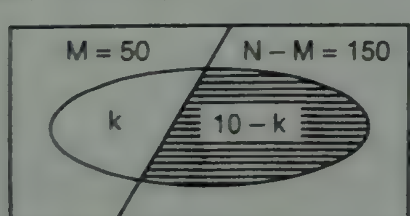
1*. A lot of N items contains M defective items. A sample of n items is drawn from the lot for quality control. Find the characteristic function of the number of defective items, contained in the sample.

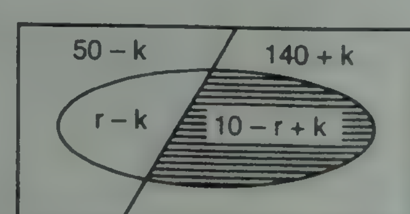
2*. A panel of 7 judges is to decide which of two final contestants A and B will be declared the winner; a sample majority of the judges will determine the winner. Assume that 4 of the judges will vote for A and the other 3 will vote for B. If we randomly select 3 of the judges and seek their verdict, what is the prob. that a majority of them will favour A ?

3*. Let X be H-G (N, M, n). Show that

$$f(x\,;N,M,n) = f(n-x\,;N,N-M,n) \qquad \text{...(1)}$$
$$= f(M-x\,;N,M,N-n) \qquad \text{...(2)}$$
$$= f(N-n-M+x\,;N,N-M,N-n) \qquad \text{...(3)}$$

4*. If $X \sim$ H-G (200, 50; 10) and $Y \mid X = k \sim$ H-G (190, 50 – k ; 10), show that $X + Y \sim$ H-G (200, 50; 20).





Exercise 15(b)

1. (a) An urn contains M red and $N-M$ white balls ($M < N$). A ball is drawn n times without replacement ($n \le N$). Find the prob. of getting x red balls among the n balls drawn. Show that the mean of distribution is nM/N [§15-22].

   (b) Show how the hypergeometric distribution arises, by giving an example. Obtain the frequency function of a random variable X following the above law. Derive $\mathrm{E}(X^{(r)})$ and Var (X).

Show that under certain conditions to be stated, the Binomial and Poisson distributions are special cases of the hypergeometric distribution.

2. (a) A receiver contains six transistors, of which two are defective. Three transistors are selected at random and inspected. Let X be the number of defectives observed. Show that the probability distribution is given by the table (ordered series) : $f(0) = 0.2, f(1) = 0.6, f(2) = 0.2$.

   (b) An urn contain w white balls and b black balls, n balls are drawn out at random, (i) with replacement, (ii) without replacement. Show that the expected number of white balls in the sample is $nw/(w + b)$.

3. There are three cartons of 12 apples each. Two of these cartons contains four rotten apples and the remaining one contains three rotten apples. You buy a carton at random and choose three apples from it. Let X be the number of rotten apples chosen. Find the probability distribution of X.

4. An urn contains n white and n green balls. We draw balls successively and without replacement one after another until all the $2n$ balls are drawn out. Find the chance that in any stage of the trial, the same number of white and green balls has been drawn.

5. In sampling a lot of 100 items, the sampling plan calls for inspection of 20 pieces. If there are any bad pieces we reject the lot, otherwise we accept it. Find the probability of accepting a lot with 5 defectives, if we allow 1 defective in the sample.

6. Establish the following recurrence relations : [Vide Example 15-5]

$$f(x\,;N, M+1, n) = \frac{(M+1)(N-M-n+x)}{(N+M)(M+1-x)}\, f(x\,;N, M, n)$$

$$f(x\,;N, M, n+1) = \frac{(N-M-n+x)(n+1)}{(n+1-x)(N-n)}\, f(x\,;N, M, n)$$

$$f(x\,;N, M+1, n) = \frac{(N+1-n)(N+1-M)}{(N+1-n-M+x)(N+1)}\, f(x\,;N, M, n).$$

7. Suppose that 100 cards marked 1, 2, ..., 100 are randomly arranged in a line. Show that the number of even integers in the first 20 positions is hypergeometrically distributed with parameters $N = 100, M = 50, n = 20$.

8. Six cards are drawn without replacement from an ordinary pack. Find the joint density of the number of aces X and the number of Queens Y. Show that the conditional density of X, given Y is

$$f(x \mid y) = {}^4C_x \,.\, {}^{44}C_{6-x-y} \,/\, {}^{48}C_{6-y}, \qquad 0 \le x+y \le 6, \quad 0 \le x, y \le 4.$$

9. Show that, as $n, N, M \to \infty$ such that $(n/N) \to 0$, $n(M/N) \to m$, the H-G distribution tends to Pois (m).

10. The k-dim H-G distribution has the p.m.f.

$$f(x_1, ..., x_k) = P(X_1 = x_1), ..., (X_k = x_k) = \binom{Np_1}{x_1} \cdots \binom{Np_{k+1}}{x_{k+1}} \Big/ \binom{N}{n}$$

where $x_{k+1} = n - (x_1 + ... + x_k)$, $p_{k+1} = 1 - (p_1 + ... + p_k)$, $x_i > 0$, $p_i > 0$.

Show that $P\{X_k \mid X_1, ..., X_{k-1}\}$ has H-G distribution.

MULTINOMIAL DISTRIBUTION

15-30. *Definition*

A random vector $X = (X_1, X_2, ..., X_k)$ possessing the probability law

$$f(x_1, x_2, ..., x_k) = \binom{n}{x_1, x_2, ..., x_k} p_1^{x_1} \cdot p_2^{x_2} ... p_k^{x_k}; 0 < p_i < 1, \forall\ i; f(x) = 0, \text{ otherwise} \quad ...(1)$$

where $x_1 + x_2 + ... + x_k = n$, $p_1 + p_2 + ... + p_k = 1$ is called a *multinomial (vector) variate* The distribution (1) is called the *multinomial distribution*. We often write "X is multin (n, p_i)", $1 \le i \le k$ to indicate that $(X_1, X_2, ..., X_k)$ has the probability law (1). The constants $n, p_1, p_2, ..., p_k$ are known as the *parameters* of multinomial distribution. A random vector X obeying law (1) is written $X \sim$ Multin $(n ; k, p_i)$ or $X \sim M(n ; k, p_i)$.

Comment. We may write (1) as a function of $k - 1$ variables by substituting $x_k = n - x_1 - x_2 - ... - x_{k-1}$. Thus

$$f(x_1, x_2, ..., x_k) = \binom{n}{x_1, ..., x_{k-1}, n - (x_1 + ... + x_{k-1})} p_1^{x_1} ... p_k^{n-(x_1 + ... + x_{k-1})} \quad ...(2)$$

However, it is convenient to retain one *redundant* variables in order to preserve the symmetry.

Physical Conditions for the Occurrence of Multi-nomial Distribution

The trials with the following properties lead to the multinomial distribution :

1. The result of each trial can be classified into one of k disjoint categories, $C_1, ..., C_k$ (say).
2. Outcomes of successive trials are independent of all other trials.
3. The prob. of element $x \in C_k$ is p_k and is the same constant for the same category for each trial.
4. The series of trials is performed a *fixed* number of times.

15-31. *Multinomial Distribution*

Let $A_1, A_2, ..., A_k$ be exhaustively disjoint events associated with a random experiment such that $P\{A_i\} = p_i$, $p_1 + p_2 + ... + p_k = 1$. If the experiment is repeated n times, the probabilities that A_1 occurs r_1 times, A_2 occurs r_2 times, ..., A_k occurs r_k times is given by

$$P_n(r_1, r_2, ..., r_k) = \frac{n!}{r_1!\ r_2!\ ....r_k!} p_1^{r_1}\ p_2^{r_2} ... p_k^{r_k}$$

where $r_1 + r_2 + ... + r_k = n$, $0 \le p_j \le 1$, $p_1 + p_2 + ... + p_k = 1$.

A_1	A_2		A_k
p_1	p_2		p_k
r_1	r_2		r_k

Proof. The r_1 trials in which A_1 occurs can be chosen in $\binom{n}{r_1}$ ways. The remaining $(n - r_1)$ trials are in store for other events. The r_2 trials in which A_2 occurs can be chosen in $\binom{n - r_1}{r_2}$ ways. Similarly, the r_3 trials in which A_3 occurs can be chosen in $\binom{n - r_1 - r_2}{r_3}$ ways and so on. Consequently, the number of ways in which events $A_1, A_2, ..., A_k$ can

happen is, by sequential counting,

$$N = \binom{n}{r_1} \cdot \binom{n-r_1}{r_2} \binom{n-r_1-r_2}{r_3} \cdots \binom{n-r_1-r_2-\ldots-r_{k-1}}{r_k} = \frac{n!}{(r_1!)(r_2!)\ldots(r_k!)}$$

Now consider any one of these ways in which events $A_1, A_2, \ldots, A_k$ occur. Since the n trials are independent, so must be trials $r_1, r_2, \ldots, r_k$. Hence

$P\{A_1$ occurs r_1 times, A_2 occurs r_2 times, ..., A_k occurs r_k times$\}$

$= P(A_1$ occurs r_1 times$)$ $P(A_2$ occurs r_2 times$)$ ... $P(A_k$ occurs r_k times$)$

$= (p_1)^{r_1} . (p_2)^{r_2} \ldots (p_k)^{r_k}$.

Since all the ways in which these events happen are disjoint,

$$\therefore \; P_n(r_1, r_2, \ldots, r_k) = N . (p_1)^{r_1} (p_2)^{r_2} \ldots (p_k)^{r_k}$$

$$= \frac{n!}{(r_1!)(r_2!)\ldots(r_k!)} (p_1)^{r_1} . (p_2)^{r_2} \ldots (p_k)^{r_k}.$$

Cor. For a Trinomial distribution

$$P(x_1, x_2, x_3) = \frac{n!}{x_1!\, x_2!\, x_3!} p_1^{x_1} p_2^{x_2} p_3^{x_3} \equiv \binom{n}{x_1, x_2, x_3} p_1^{x_1} p_2^{x_2} p_3^{x_3}$$

where $\quad x_1 + x_2 + x_3 = n, \; 0 \le p_i \le 1, \; p_1 + p_2 + p_3 = 1.$

15-32. M.G.F. of Multinomial Distribution

$$M(t_1, t_2, \ldots, t_k) = (p_1 e^{t_1} + p_2 e^{t_2} + \ldots + p_k e^{t_k})^n \qquad [\S\ 8\text{-}16\ (6)] \qquad \ldots(1)$$

1. *Marginal Distributions of Multinomial Law :*

To find p.m.f. of X_i put $t_j = 0$, $(j \ne i)$, $j = 1, 2, \ldots, k$, then

$$M(t_1 : X_i) = M_X(0, \ldots, 0, t_i, 0, \ldots 0) = (p_i e^{t_i} + \Sigma p_j)^n = (p_i e^{t_i} + 1 - p_i)^n.$$

$$\therefore \quad M(t_i : X_i) = (q_i + p_i e^{t_i})^n. \qquad [q_i = 1 - p_i]. \qquad \ldots(2)$$

Thus $X_i \sim$ bin (n, p_i), i.e. each marginal (individual) variate is binomial.

2. *Trinomial Distribution :*

The joint distribution of $(X_i, X_j, n - X_i - X_j)$ is again multin$(n; p_i, p_j, 1 - p_i - p_j)$.

Proof. $\quad M(t_1, t_2, t_3) = \mathrm{E}\{\exp[t_1 X_1 + t_2 X_2 + t_3(n - X_1 - X_2)\}$. [$i = 1, j = 2$, for simplicity]

$$= e^{nt_3} \mathrm{E}\{e^{(t_1 - t_3)X_1 + (t_2 - t_3)X_2}\} = e^{nt_3}(p_1 e^{t_1 - t_3} + p_2 e^{t_2 - t_3} + p_3 + \ldots + p_n)^n. \qquad \text{[by (1)]}$$

$$= [p_1 e^{t_1} + p_2 e^{t_2} + (1 - p_1 - p_2) e^{t_3}]^n \qquad (\because \Sigma p_k = 1), \qquad \ldots(3)$$

This completes the proof. Thus

$$f(x_i, x_j) = \{n!/\, x_i!\, x_j!\, (n - x_i - x_j)!\}\, p_i^{x_i} . p_j^{x_j} (1 - p_i - p_j)^{n - x_i - x_j}. \qquad \ldots(4)$$

3. *Conditional Distribution in Trinomial Law* : $\{X_1 \mid X_2 = v\} \sim \text{bin}\,[n - v, p_1/(1 - p_2)]$.

Proof. $P\{X_1 = u \mid X_2 = v\} = P\{X_1 = u, X_2 = v\} / P\{X_2 = v\}$ (by Def.)

$$= P\{X_1 = u, X_2 = v, n - X_1 - X_2 = n - u - v\} / P\{X_2 = v\}$$

$$\text{Numerator} = \frac{n!}{u!\,v!\,(n-u-v)!} p_1^u p_2^v (1 - p_1 - p_2)^{n-u+v} \qquad \text{[by (3)]}$$

$$\text{Denominator} = \frac{n!}{v!\,(n-v)!} p_2^v (1 - p_2)^{n-v}, \quad [X_2 \sim \text{bin}\,(n, p_2)].$$

$$\therefore \quad P\{X_1 = u \mid X_2 = v\} = \binom{n-v}{u}\left(\frac{p_i}{1-p_2}\right)^u \left(1 - \frac{p_1}{1-p_2}\right)^{n-uv} = \binom{n-v}{u} q^{n-u-v} p^u.$$

15-33. *Sum of Individual Components of Multinomial Distribution*

$$M(t : X_i + X_j) = \mathsf{E}\{e^{t(X_i + X_j)}\} = M(0, ..., 0, t_i, 0, 0, t_j, 0, ..., 0), \qquad (t_i = t_j = t)$$

$$= (p_1 + ... + p_i e^t + p_{i+1} + ... + p_j e^t + p_{j+1} + ... + p_k)^n$$

$$= [1 - p_i - p_j + (p_i + p_j)e^t] = (q + pe^t)^n, \quad [p = p_i + p_j, q = 1 - p]$$

This shows that $(X_i + X_j) \sim \text{bin}\,(n, p_i + p_j)$.

Note. $(X_i + X_j + X_k) \sim \text{bin}\,(n, p_i + p_j + p_k)$, $i \neq j \neq k$ and so on.

15-34. *Correlation Coefficient*

Since X_i is bin (n, p_i), we have Var $(X_i) = np_i\, q_i$.

Since $X_i + X_j$ is bin $(n, p_i + p_j)$, Var $(X_i + X_j) = n(p_i + p_j)(1 - p_i - p_j)$.

$$\because \quad \text{Var}\,(X_i + X_j) = \text{Var}\,(X_i) + \text{Var}\,(X_j) + 2\,\text{Cov}\,(X_i, X_j)$$

$$\therefore \quad 2\,\text{Cov}\,(X_i, X_j) = [n(p_i + p_j) - n(p_i + p_j)^2] - np_i\, q_i - np_j\, q_j = -2np_i\, p_j$$

$$\rho(X_i, X_j) = \sigma_{ij} / \sigma_i \sigma_j = -[p_i\, p_j / q_i\, q_j]^{1/2}. \qquad \text{[independent of } n\text{]} \quad ...(i)$$

***Note* : 1.** $M(t_i, t_j) = M(0, ..., 0, t_i, 0, ... 0, t_j, 0, ..., 0) = (p_i e^{t_i} + p_j e^{t_j} + p_1 + ... p_k)^n$.

$$\mathsf{E}(X_i\, X_j) = \left(\frac{\partial^2 M(t_i, t_j)}{\partial t_i\, \partial t_j}\right)_{\vec{t} = 0} = n(n-1)\, p_i\, p_j [e^{t_i + t_j}(p_i e^{t_i} + p_j e^{t_j} + p_1 + ... + p_k)^{n-2}]_{t=0}$$

$$= n(n-1)\, p_i\, p_j, \qquad (i \neq j).$$

$$\text{Cov}\,(X_i, X_j) = \mathsf{E}(X_i\, X_j) - \mathsf{E}(X_i)\mathsf{E}(X_j) = n(n-1)\, p_i\, p_j - n^2 p_i\, p_j = -np_i\, p_j.$$

2. Observe that $S_k = X_1 + ... + X_k$ = total sum = Const. is degenerate, since Var (const) = 0

$$0 = \text{Var}\,(X_1 + ... + X_k) \neq \text{Var}\,(X_1) + ... + \text{Var}\,(X_k). \qquad [X_j \text{ are dependent}]$$

15-35. Reproductive Property

If $X \sim$ Multin $(m\ ;\ p_i)$ and $Y \sim$ Multin $(n\ ;\ p_i)$ $1 \le i \le k$, are independent random vectors, then $X + Y$ is Multin $(m + n\ ;\ p_i)$.

Proof. Since X and Y are *independent* k-component random variables

$$M(t : X + Y) = M(t : X) \,.\, M(t : Y) = (\Sigma p_i\, e^{t_i})^m\, (\Sigma p_i\, e^{t_i})^n = (\Sigma p_i\, e^{t_i})^{m+n}.$$

Thus $X + Y \sim$ Multin $(m + n, p_i)$, $1 \le i \le k$, by uniqueness theorem of m.g.f.'s.

Remark. This property characterizes the multinomial distribution.

15-36. Worked-out Problems

Example **1.** The trinomial distribution of two variates X, Y is defined by

$$f_{X,Y}(x, y) = \frac{n!}{x!\, y!\, (n - x - y)!}\, p^x q^y (1 - p - q)^{n-x-y},$$

for $x, y = 0, 1, \ldots, n$ and $x + y \le n$, where $0 \le p,\ q \le 1,\quad p + q \le 1$.

(i) Find the p.g.f. of X, Y, Z, $[Z = n - X - Y]$.

(ii) Show that Var $(X + Y + Z) \ne$ Var $(X + Y + Z)$.

(iii) Find $f_X(x), f_Y(y), f_{X \mid Y}(x \mid y)$, $\mathsf{E}(X \mid Y)$ and Corr (X, Y).

Solution. (i) We rewrite : $p = p_1$, $q = p_2$, $1 - p - q = p_3$, $n - x - y = z$, so that

$$\mathsf{E}(t_1^X\, t_2^Y\, t_3^Z) = \sum_{x,y,z} \binom{n}{x, y, z} p_1^x\, p_2^y\, p_3^z\, t_1^x\, t_2^y\, t_3^z = \sum_{x,y,z} \binom{n}{x, y, z} (p_1 t_1)^x\, (p_2 t_2)^y\, (p_3 t_3)^z.$$

$$\therefore\ G(t_1, t_2, t_3) = \mathsf{E}(t_1^X \,.\, t_2^Y \,.\, t_3^Z) = (p_1 t_1 + p_2 t_2 + p_3 t_3)^n. \quad \text{[By Multinomial Expansion]}$$

(ii) Since $X + Y + Z = n$ (fixed), Var (Const) = 0, so

$\therefore\ 0 = \text{Var}\,(X + Y + Z) \ne \text{Var}\,(X) + \text{Var}\,(Y) + \text{Var}\,(Z)$, $\Rightarrow$ $X + Y + Z$ is degenerate.

To get rid of *redundant* variate Z, we put $t_3 = 1$ and the p.g.f. of X and Y is thus

$$G(t_1, t_2) = [pt_1 + qt_2 + (1 - p - q)]^n \equiv G(t_1, t_2, 1) \qquad \ldots(1)$$

(ii) $$G(t, 1, 1) = [(1 - p) + pt]^n \Rightarrow X \sim \text{bin}\,(n, p).$$

$$G(1, t, 1) = [(1 - q) + qt]^n \Rightarrow Y \sim \text{bin}\,(n, q).$$

Noting $G(e^t : 1) = M(t : X)$ provides $M(t_1, t_2) \ne M(t_1)\, M(t_2)$, so X and Y are not independent.

$$P(X = x \mid Y = y) = \frac{f(x, y)}{f_Y(y)} = \left[\frac{n!\, p^x\, q^y\, (1 - p - q)^{n-x-y}}{x!\, y!\, (n - x - y)!}\right]\left(\frac{(n - y)!\, y!}{n!\, q^y\, (1 - q)^{n-y}}\right)$$

$$= \binom{n - y}{x}\left(\frac{p}{1 - q}\right)^x \left(1 - \frac{p}{1 - q}\right)^{n-x-y}, \quad x = 0, 1, 2, \ldots, n - y.$$

Thus, $\quad (X \mid Y = y) \sim \text{bin}\,[n - y, p/(1 - q)];\quad 0 < p/(1 - q) \le 1.$

Similarly, $(X \mid Y = x) = \sim \text{bin}\,[n - x, q/(1 - p)]$. It follows that

$$\mathsf{E}(X \mid y) = (n - y)p/(1 - q) \quad \mathsf{E}(Y \mid x) = (n - x)\,q/(1 - p) \qquad \ldots(2)$$

Now : $$\mathrm{E}(XY) = \frac{\partial^2 G(1,1)}{\partial t_1\, \partial t_2} = n(n-1)\, pq\, [p.1 + q.1 + (1-p-q)]^{n-2} = n(n-1)\, pq.$$

$$\mathrm{Cov}\,(X, Y) = \mathrm{E}(XY) - \mathrm{E}(X)\,\mathrm{E}(Y) = (n^2 pq - npq) - (np)\,(nq) = -\,npq.$$

$$\rho(X, Y) = \frac{\sigma_{XY}}{\sigma_X\, \sigma_Y} = \frac{-npq}{\sqrt{np\,(1-p)\,.\,nq\,(1-q)}} = -\left[\frac{pq}{(1-p)\,(1-q)}\right]^{1/2} \qquad \ldots(3)$$

Note. From the known properties of linear regression coefficients, using (2)

$$\rho^2 = (\text{Coeff. of } x)\ (\text{Coeff. of } y) = pq/(1-p)\ (1-q).$$

Thus ρ is given by ($\rho < 0$) since regression coefficients are negative.

Example 2. A trial can result into one of three possible outcomes A_1, A_2, A_3 with probability p_1, p_2, p_3 ; $p_1 + p_2 + p_3 = 1$. This trial is performed n independent times and outcomes A_j occurs X_j times (j = 1, 2, 3). Determine the conditional expectation of X_1 given $X_2 = m$.

Solution. For $X_1 = k$, $X_2 = m$, we automatically get $X_3 = n - m - k$. Hence

$$P\,\{X_1 = k \mid X_2 = m\} = \frac{P\{X_1 = k, X_2 = m\}}{P\{X_2 = m\}} \equiv \frac{P\{X_1 = k, X_2 = m, X_3 = n - m - k\}}{P(X_2 = m)}$$

$$= \left[\frac{n!\, p_1^k \cdot p_2^m \cdot p_3^{n-m-k}}{k!\, m!\, (n-m-k)!}\right] \Big/ \left[\frac{n!}{m!\,(n-m)!} \cdot p_2^m\, q_2^{n-m}\right], \quad [X_2 \sim \text{bin}\,(n, p_2)]$$

$$= \frac{(n-m)!}{(n-m-k)!\,k!}\left(\frac{p_1}{q_2}\right)^k \cdot \left(\frac{p_3}{q_2}\right)^{n-m-k} \quad \left[\frac{p_3}{q_2} = \frac{1 - p_2 - p_1}{q_2} = 1 - \frac{p_1}{q_2}\right]$$

$$= \binom{n-m}{k}\left(\frac{p_1}{q_2}\right)^k \left(1 - \frac{p_1}{q_2}\right)^{n-m-k}.$$

This shows that $(X_1 = k \mid X_2 = m) \sim \text{bin}\,(n - m, p_1 / q_2)$

This provides : $\mathrm{E}\,(X_1 \mid X_2 = m) = (n-m)\,(p_1 / q_2)$. $\qquad (q_2 = 1 - p_2)$

Example 3. Independent variates X_i, $1 \le i \le n$ are Poisson with parameters $\lambda_i, 1 \le i \le k$. Let $S = X_1 + X_2 + \ldots + X_k$. Prove that the conditional distribution $P\{X_1, X_2, \ldots, X_k \mid S = n\}$, is multinomial.

Solution. Recalling the reproductive property of the Poisson variates, we have $S \sim \text{pois}\,(\Sigma\lambda_j)$. Now

$$p = P\{X_1, X_2, \ldots, X_k \mid S = n\} = P\{X_1 = r_1, X_2 = r_2, \ldots, X_k = n - r_1 \ldots - r_{k-1}\} / P(S = n)$$

$$= P(X = r_1) \ldots P(X_{k-1} = r_{k-1})\,.\,P(X_k = n - r_1 \ldots - r_{k-1}) / P(S = n) \qquad [\because X_i \text{ are indep.}]$$

$$= \frac{e^{-\lambda_1}(\lambda_1)^{r_1}}{r_1!} \cdots \frac{e^{-\lambda_{k-1}}(\lambda_{k-1})^{r_{k-1}}}{r_{k-1}!} \cdot \frac{e^{-\lambda_k}(\lambda_k)^{r-r_1 \ldots r_{k-1}}}{(n - r_1 \ldots - r_{k-1})!} \Big/ \frac{e^{-\lambda}(\lambda)^n}{n!}, [\lambda = \Sigma\lambda_j]$$

$$= \frac{n!}{r_1! \ldots r_{k-1}!\, r_k!}\left(\frac{\lambda_1}{\lambda}\right)^{r_1} \cdots \left(\frac{\lambda_k}{\lambda}\right)^{r_k}, \quad p_j = \frac{\lambda_j}{\lambda}.$$

Obviously $\Sigma r_j = n$, $\Sigma p_j = 1$. We have established that

$$(X_1 \cap X_2 \dots \cap X_k) \mid \Sigma X_j = n \text{ is multin } (n\ ;\ p_1, p_2, \dots p_k).$$

Remarks. If $\lambda_i = m$, $\forall\ i$, then $p_i = 1/k$, so that probability in multinomial set is $1/k$ (constant).

Problems with Solutions Provided at the End of the Text

1*. Three fair dice are cast. In 10 independent casts, let X be the number of times all three faces are alike and Y be the number of times only two faces are alike. Find the joint p.d.f. of X and Y and compute $\mathsf{E}(24\,XY)$.

2*. For a Trinomial distribution [§15-36, Example 1] find the m.g.f. of the variate $Z = n - X - Y$. Hence or otherwise find Var (Z) and Cov (Y, Z).

3*. Prove that the multinomial distribution (§15-31) can be approximated by the multi-dimensional Poisson law :

$$e^{-(\lambda_1 + \dots \lambda_n)} (\lambda_1)^{k_1} (\lambda_2)^{k_2} \dots (\lambda_m)^{k_m} / k_1!\, k_2! \dots k_m!$$

p_{m+1} are small and n is large but $np_j = \lambda_j$ fixed.

4*. Show that the $k \times k$ covariance matrix of a $M(n\ ;\ p_1, p_2, \dots, p_k)$ distribution is singular with rank $k - 1$.

5*. Of the 100 people in a certain village 50 always tell the truth, 30 always lie and 20 always refuse to answer. A single unbiased die is tossed. If the result is 1, 2, 3 or 4, (event A) a sample of size 30 is taken *with replacement*. If the result is 5 or 6 (event B) a sample of size 30 is taken *without replacement*. Find $\mathsf{E}(X)$ when X is defined by

$X = 1$, if resulting sample contains 10 people of each category.

$X = 2$, if the sample is taken with replacement and contains 12 liars; $X = 3$, otherwise.

Exercise 15(c)

1. (a) Team A is going to play five games with team B. It is known that team A wins any match with team B with probability 0.5, while team B wins with probability 0.3 [with the remaining games being tied]. What is the probability that team A will win two games and lose one ? Consider successive games independent.

   (b) Assume that team A can beat team B with probability 0.4 while team B will defeat team A with probability 0.3 and three is also some chance of a tie. What is the probability that team A will win a play-off series consisting of four games ?

2. (a) Two dice are cast 10 times. Let X be the number of times no 1s appear and let Y be the number of times two 1s appear. Find $P\{X < 3, Y < 3\}$, $P\{X + Y = 4\}$ and $P\{2 \le X + Y \le 4\}$.

   (b) A die is cast 12 times. Find the chance that every face turns up twice. If X and Y denote the frequency of turning up of 5 and 6, find the joint distribution of X, Y and show that Cov $(X, Y) = -1/3$.

3. A and B play 12 games of chess of which 6 are won by A, 4 by B and the remaining games being tied. In a tournament of 3 games, find the probability (i) A wins 3 games (ii) B wins at least one game.

4. (a) A box contains 5 red, 4 white and 3 black balls. A ball is selected at random from the box, its colour noted and then the ball is replaced. Find the chance that out of 6 balls selected in this manner, 3 are white, 2 are red and 1 is black.

(b) A box contains equal number of white, blue, yellow and pink balls. Four balls are drawn independently and with replacement. Find the probability that exactly k, $(1 \le k \le 4)$ colours will appear in the sample.

5. A group consists of 50% Indians, 30% Americans and 20% Britishers. If a sample of 6 persons is selected at random, find the chance that 2 are Indians, 3 are Americans and 1 is a Britisher.

6. (a) Three coins are tossed n times. Let X, Y, Z denote respectively the number of times no head, 1 head, 2 head appear. Find the joint density $f(x, y, z)$ as well as $f(x, z \mid y)$.

(b) Two events A and B associated with an experiment, have respectively probabilities of occurrence p and p_1. Show that in n trials, the probability that AB occurs r_1 times ; AB' : r_2 times ; $A'B$: r_3 times, $A'B'$: r_4 times is

$$\{n!/r_1!r_2!r_3!r_4!\}.\,p^{r_1+r2}\,p_1^{r_3+r_4}(1-p)^{r_3+r_4}(1-p_1)^{r_3+r_4}.$$

7. Show that the mode (most probable value) of the multin (n, k, p_i) is given by $x_1, x_2, ..., x_k$, satisfying $np_i - 1 < x_i \le (n+k-1)\,p_i$, $i = 1, 2, ..., k$.

8. Suppose $(Y_1, Y_2, Y_3, Y_4, Y_5)$ is multinomially distributed with parameters $n, p_1, p_2, p_3, p_4, p_5$. What is the distribution of $(Y_1 + Y_2, Y_3, Y_4 + Y_5)$.

9. Suppose (Y_1, Y_2, Y_3) is multin $(n; p_1, p_2, p_3)$. Show that the conditional distribution of Y_1, given that $Y_1 + Y_2 = r$, is bin $[r, p_1/(p_1, p_2)]$.

10. Let (Y_1, Y_2, Y_3, Y_4) be multin (n, p, p_2, p_3, p_4). Prove that the conditional distribution of (Y_1, Y_2, Y_3) given $Y_4 = k$ is multin $[n-k, p_1/(1-p_4), p_2/(1-p_4), p_3/(1-p_4)]$.

11. If $X = (X_1, X_2, ..., X_k)$ is multin $(n, p_1, p_2, ..., p_k)$ find $\mathrm{E}(Y)$ and Var (Y), where

$$Y = \Sigma\{(X_i - np_i)^2/np_i\}, i = 1, 2, ..., k.$$

12. A chain smoker has $(n+1)$ match boxes $B_1, ..., B_{n+1}$ each containing N matches. Any time he lights a cigarette, he chooses one of the boxes at random. Let $X_1, ..., X_n$ be the numbers of matches left in the remaining n boxes. Find the chance when (i) a box is found empty for the first time (ii) Any box is emptied first.

Suppose that B_k, $(1 \le k \le n)$ is chosen with prob. p_k. Find the chance that B_k empties first.

Note. The case $n = 1$ is called Banach match-box problem.

15-40. *Logarithmic Distribution*

Since, $-\ln(1-q) = \Sigma(q^x/x)$, $1 \le x < \infty$, we multiply it by $(-\ln p)^{-1}$ to obtain $(-\ln p)^{-1}\,\Sigma(q^x/x) = 1$, which motivates.

Definition. A r.v. X is said to possess logarithmic distribution with parameter p, if its p.m.f. is

$$P\{X = x\} = f(x) = Lq^x/x, \quad L = (-\ln p)^{-1} \quad x = 1, 2, 3, ... \qquad ...(1)$$

15-41. *Moment Generating Function and Moments*

$$M(t : X) = \mathrm{E}(e^{tX}) = L\Sigma(e^{tx}q^x/x) = L\Sigma(qe^t)^x/x = -L\ln(1-qe^t), \quad x = 1, 2, ... \qquad ...(2)$$

Assuming $|qe^t| < 1$, we expand (3) to get

$$M(t : X) = L\sum_{r=1}^{\infty}\frac{q^r e^{rt}}{r} = \sum_{n=0}^{\infty}\sum_{r=1}^{\infty} Lq^r.r^{n-1}.\frac{t^n}{n!} \qquad ...(3)$$

$$\therefore \quad \mu'_n = \mathsf{E}(X^n) = L \sum_{r=1}^{\infty} q^r r^{n-1}, \quad n = 1, 2, 3, \ldots \qquad \ldots(4)$$

Thus $\mathsf{E}(X) = L\Sigma q^r = Lq/(1-q) = Lq/p$, $\mathsf{E}(X^2) = L\Sigma q^r r = Lq/(1-q)^2 = Lq/p^2$

$\mathsf{E}(X^3) = L\Sigma q^r r^2 = Lq(1+q)/(1-q)^3 = Lq(1+q)/p^3$.

$\mathsf{E}(X^4) = L\Sigma q^r r^3 = L[2q^2 + q(1+q)^2]/(1-q)^4 = Lq[2q + (1+q)^2]/p^4$.

Var $(X) = Lq/p^2 - (Lq/p)^2 = (Lq/p^2)[1 - Lq]$.

$$\mu_3 = Lq(1+q)/p^3 - 3(Lq/p^2).(Lq/p) + 2(Lq/p)^3 = Lq\{q + (1-Lq)(1-2Lq)\}/p^3.$$

$$\begin{aligned} \mu_4 &= \mu'_4 - 4\mu'_3 \mu + 6\mu'_2 \mu'^2_1 - 3\mu'^4_1 \\ &= Lqp^{-4}\{2q + (1+q)^2 - 4Lq(1+q) + 6L^2q^2 - 3L^3q^3\} \\ &= Lqp^{-4}\{1 + 4q + q^2 - 4Lq - 4Lq^2 + 6Lq^2 - 3L^3q^3\}. \end{aligned}$$

$$\therefore \quad \gamma_1 = \sqrt{\beta_1} = \mu_3/(\mu_2)^{3/2} = (Lq)^{-1/2}(1-Lq)^{-3/2}\{q + (1-Lq)(1-2Lq)\}.$$

$$\beta_2 = (\mu_4/\mu_2^2) = (1-Lq)^{-2}\{1 + 4q + q^2 - 4Lq - 4Lq^2 + 6L^2q^2 - 3L^3q^3\}(Lq)^{-1}$$

15-42. *Decreasing Factorial Moments*

$$\mu_{(r)} = \mathsf{E}[X^{(r)}] = L\sum_{x=1}^{\infty} x^{(r)}\frac{q^x}{x} = Lq^r \sum_{x=1}^{\infty}\frac{x!}{(x-r)!}\frac{q^{x-r}}{x} = Lq^r\sum_{x=r}^{\infty}\frac{(x-1)!}{(x-r)!}q^{x-r}.$$

$$= Lq^r\sum_{k=0}^{\infty}\frac{(k+r-1)!}{k!}q^k = (r-1)!\,Lq^r\sum_{k=0}^{\infty}\binom{k+r-1}{k}q^k, [x-r=k]$$

$$= Lq^r(r-1)!(1-q)^{-r} = L(q/p)^r.(r-1)!. \qquad \left[1-(x)^{-n} = \sum_{k=0}^{\infty}\binom{n+k-1}{k}x^k\right]$$

This provides, $\mathsf{E}(X) = L(q/p)$, $\mathsf{E}\{X(X-1)\} = L(q/p)^2 \Rightarrow \mathsf{E}(X^2) = Lq/p^2$.

$\mathsf{E}[X(X-1)(X-2)] = 2L(q/p)^3 \Rightarrow \mathsf{E}(X^3) = Lq(1+q)/p^3$, etc.

Example : Show that under suitable conditions (to be stated) the logarithmic-series distribution may be obtained as a limiting case of the N-B distribution.

Solution. If X is neg-bin (k, p) and Y is ℓn-series distributed with parameter α, then

$$f(x) = \binom{x+k-1}{k-1}p^k q^x, \quad x = 0, 1, 2, 3, \ldots \qquad \ldots(1)$$

$$g(y) = [1 - \ell n(1-\alpha)]^{-1}\alpha^y/y, \quad y = 1, 2, 3, \ldots \qquad \ldots(2)$$

Since the range of ℓn-series distribution starts from $y = 1$, we truncate neg-bin distribution at $x = 0$. Consequently, we divide (1) by $P(X \geq 1) = 1 - p^k$, to get the truncated N-B density f_0.

$$f_0(x) = \binom{x+k-1}{k-1}\frac{p^k q^x}{1-p^k} = \frac{(x+k-1)!}{x!\,k!}\left(\frac{kp^k}{1-p^k}\right)q^x, \quad x = 1, 2, 3, \ldots \qquad \ldots(3)$$

We now let $k \to 0$ and use L'Hospital rule, to get from (3)

$$\lim_{k\to 0}\left(\frac{kp^k}{1-p^k}\right) = \lim_{k\to 0}\left(\frac{kp^k \ln p + p^k}{-p^k \ln p}\right) = \lim_{k\to 0}\left(k - \frac{1}{\ln p}\right) = \frac{-1}{\ln p}.$$

$$\therefore \qquad \lim_{k\to 0} f_0(x) = (-\ln p)^{-1}\, q^x / x \qquad x = 1, 2, 3, \ldots \qquad \ldots(4)$$

The limiting density is the same as that given by (2), with $\alpha = q$. Thus, the truncated Neg-bin distribution reduces to ℓn-series distribution, when the parameter $k \to 0$.

15-50. *Generalized Power Series Distribution (GPSD)*

Let X be a non-negative integer-valued r.v. with generating function

$$g(\theta) = \Sigma a_x\, \theta^x, \quad x = 0, 1, 2, \ldots, a_x \geq 0, \ \theta \geq 0. \qquad \ldots(1)$$

so that $0 < g(\theta) < \infty$. We assume that (1) is uniformly convergent.

Definition. A r.v. X is said to possess GPSD if its p.m.f. is

$$\boxed{f(x) = P(X = x) = a_x\, \theta^x / g(\theta), \quad a_x \geq 0, \ x = 0, 1, \ldots} \qquad \ldots(2)$$

$$\therefore \qquad M(t : X) = E(e^{tX}) = \sum_{x=0}^{\infty} \frac{a_x\, \theta^x\, e^{tx}}{g(\theta)} = \frac{g(\theta e^t)}{g(\theta)}. \qquad \ldots(3)$$

15-51. *Special Cases*

The definition (2) embodies several discrete distribution, e.g. Binomial, Poisson, Negative Bin; ℓn-series distribution, etc. We illustrate by some derivations. This shall also point out the method of attack.

1. ***Binomial distribution.*** Let $g(\theta) = (1 + \theta)^n$ with $\theta = p/q$.

Then $g(\theta) = \Sigma a_x \theta^x$ provides $(1+\theta)^n = \Sigma\, {}^nC_x\, \theta^x = \Sigma a_x\, \theta^x \Rightarrow a_x = {}^nC_x$.

$$\therefore \qquad f(x) = \binom{n}{x}\left(\frac{p}{q}\right)^x \Big/ \left(1 + \frac{p}{q}\right)^n = \binom{n}{x} p^{n-x} p^x, 0 \leq x \leq n.$$

2. ***Poisson distribution.*** Let $g(\theta) = e^0$. Then $e^\theta = \Sigma\theta^x / x! = \Sigma a_x\, \theta^x \Rightarrow a_x = 1/x!$

$$\therefore \qquad f(x) = \theta^x e^{-\theta} / x! \quad 0 \leq x < \infty.$$

3. ***Neg-bin distribution.*** Let $g(\theta) = (1 - \theta)^{-k}$, with $\theta = q$.

Then $g(\theta) = \Sigma a_x \theta^x \Rightarrow \Sigma\binom{k+x-1}{x}\theta^x = \Sigma a_x \theta^x \Rightarrow a_x = \binom{k+x-1}{x} \Rightarrow f(x) = \binom{k+x-1}{x} p^k q^x.$

4. ***Log-Series distribution.*** Let $g(\theta) = -\ell n\,(1 - \theta)$, then

$\Sigma a_x \theta^x = -\ell n(1-\theta) \Rightarrow a_x = 1/x$, hence $f(x) = \theta^x / x\,[-\ell n\,(1-\theta)]$, $1 < \theta < 0$.

15-52. Cumulant Recurrence Formula

From $K(t) = \ln N(t)$, we get from 15-50(3) $K(t) = \ln g(\theta e^t) - \ln g(\theta)$.

Replace $K(t)$ by its defining series ; differentiate it partially w.r.t. parameters θ and t to get

$$\Sigma \frac{t^r}{r!}\frac{\partial k_r}{\partial \theta} = \frac{e^t g'(\theta e^t)}{g(\theta e^t)} - \frac{g'(\theta)}{g(\theta)}, \quad \Sigma \frac{t^{r-1}}{(r-1)!} k_r = \frac{\theta e^t g'(\theta e^t)}{g(\theta e^t)}.$$

Eliminating $g'(\theta e^t)$ provides

$$\theta \Sigma \frac{t^r}{r!}\frac{\partial k_r}{\partial \theta} = \Sigma k_{r+1}\frac{t^r}{r+1} - \frac{\theta g'(\theta)}{g(\theta)}$$

Comparing various powers of t, we get

$$k_1 = \theta g'(\theta) / g(\theta), \qquad k_{r+1} = \theta\,(dk_r/d\theta).$$

Exercise. For the ln-series distribution with p.d.f.

$$f(x) = -q^x / (x \ln p), \quad x = 1, 2, \ldots, ; 0 < p < 1; q = 1 - p$$

prove the following relations :

(a) $\mu'_{r+1} = q(\partial \mu'_r / \partial q) + L(q/p)\,\mu'_r$. (b) $\mu_{r+1} = q(\partial u_r / \partial q) + r\mu_2\,\mu_{r+1}$.

(c) $\mu'_{(r)} = L\theta^r\,(r-1)!\,p^r$. (d) $k_{r+1} = \theta\,(\partial k_r / \partial \theta)$.

Success always comes at a price.

****

People who don't count won't count. (Anatole France)

Normal (Gaussian) Distribution

16

16-00. *A Special Form of a Function*

Consider $I = \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$. Put $(x - \mu)/\sigma = z$, then

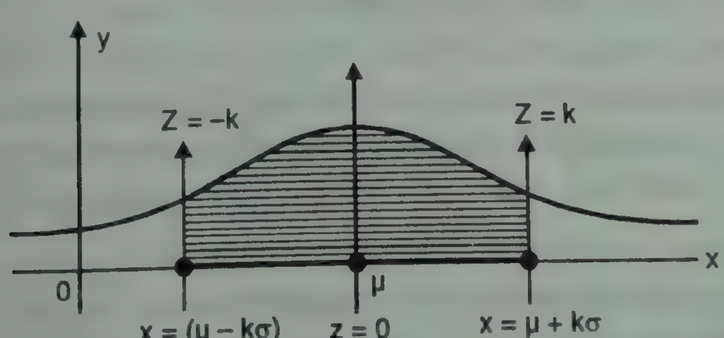
$$I = \sigma\int_{-\infty}^{\infty} e^{-\frac{1}{2}z^2} dz = 2\sigma\int_0^{\infty} e^{-\frac{1}{2}z^2} dz = 2\sigma\int_0^{\infty} e^{-t} \frac{1}{\sqrt{2}} t^{\frac{1}{2}-1} dt \quad \left(\frac{1}{2}z^2 = t\right)$$

$$= \sqrt{2}\sigma \,.\, \Gamma\left(\tfrac{1}{2}\right) = \sigma\sqrt{2\pi}.$$

Conclusion : $\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} . e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx = 1$. Thus, we introduce :

Definition. A r.v. X is said to possess Normal (or Gaussian) distribution, with parameters μ and σ^2, notated $N(\mu, \sigma^2)$ if its p.d.f. is

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2} = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \; -\infty < x < \infty . \qquad \text{...(1)}$$



Warning. Some authors write $N(\mu, \sigma)$ instead of $N(\mu, \sigma^2)$. We shall insist on $N(\mu, \sigma^2)$.

Distribution Function : $\Phi(x) = P(X \le x) = \int_{-\infty}^{\infty} f(x)\, dx$ ...(2)

$\therefore \quad P\{a < X \le b\} = \Phi(b) - \Phi(a) = \int_a^b f(x)\, dx.$ ...(3)

The integral in (2) cannot be evaluated by elementary techniques. Its Tabulation is provided at the end of this book.

***Example* 1.** Two r.v.s with the same distribution function may not be identical. Can X and $-X$ have the same distribution ? If so, when ?

Solution. Let $X \sim N(0, 1)$ and take $Y = -X$. Then

$P(Y \le y) = P(-X \le y) = P(X \ge -y) = (X \le y)$, by **symmetry.** [Draw a figure.]

***Example* 2.** If $f(x) = Ce^{-3x^2+6x}$, $-\infty < x < \infty$, evaluate C, μ, σ^2.

Solution. Here $f(x) = Ce^3 . e^{-3(x-1)^2} = Ce^3 \exp[-(x-1)^2/2(1/6)]$. [Gaussian Form]

Compare it with $f(x) = (\sigma\sqrt{2\pi})^{-1} \exp[-(x-\mu)^2/2\sigma^2]$, to get $\mu = 1$, $\sigma^2 = 1/6$, $C = (3/\pi e^6)^{1/2}$.

16-12. *Standardized (or normalized) Normal Variate*

Setting $z = (x-\mu)/\sigma$, reduces $N(\mu, \sigma^2)$ to $N(0, 1)$ which is called standard (or unit) normal distribution. The c.d.f. of $N(0, 1)$ is thus

$$\Phi(z) = P(Z \le z) = \int_{-\infty}^{z} f(z)\,dz, \qquad \left[f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}\right]. \qquad \text{...(1)}$$

$\Phi(z)$ has been thoroughly tabulated. We also frequently use

$$\boxed{\psi(k) = \int_0^k f(z)\,dz = \Phi(k) - 0.5} \qquad \text{...(2)}$$

$$\Phi(k) = \int_{-\infty}^{0} f(z)\,dz + \int_0^k f(z)\,dz = 0.5 + \psi(k). \qquad \text{...(3)}$$

Standard Valuations. $\psi(1) = 0.3413$, $\psi(2) = 0.4772$, $\psi(3) = 0.4987$
$\psi(1.96) = 0.4750$, $\psi(2.58) = 0.4951$, $\psi(1.645) = 0.4500$.

16-20. *Tabulation of Normal Distribution*

$$P\{a \le X \le b\} = \int_a^b f(x)\,dx = \int_\mu^b f(x)\,dx - \int_\mu^a f(x)\,dx = \left\{\int_0^{(b-\mu)/\sigma} - \int_0^{(a-\mu)/\sigma}\right\} f(z)\,dz, \left[z = \frac{x-\mu}{\sigma}\right]$$

$$= \psi[(b-\mu)/\sigma] - \psi[(a-\mu)/\sigma] \text{ where } f(z) = (\sqrt{2\pi})^{-1} e^{-z^2/2}. \text{ [p.d.f. of } N(0,1)]$$

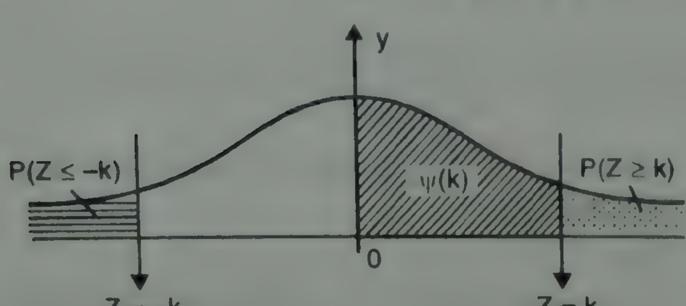
We thus notice that the evaluation of $P\{a \le X \le b\}$ is related to the integral $\psi(k)$ which represents area bounded by the curve $y = f(z)$ and the ordinates $z = 0$ and $z = k$. The definite integral $\psi(k)$, also called *normal probability integral*, cannot be evaluated by elementary techniques and has been tabulated numerically for different values of k at intervals of 0.01.

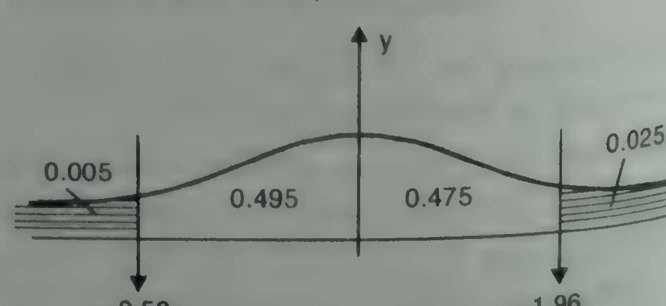
Another table related to the ordinates of standard normal curve $y = (\sqrt{2\pi})^{-1} \exp(-\frac{1}{2}z^2)$ is also presented in the Appendix.

16-21. *Special Limits*

Let $X \sim N(\mu, \sigma^2)$ put $z = (x-\mu)/\sigma$ i.e. $x = \mu + \sigma z$; then

$$P\{\mu - k\sigma \le X \le \mu + k\sigma\} = P\{-k \le Z \le k\} = 2\psi(k).$$





Using tables of $N(0, 1)$-area we obtain

$P\{|Z| \leq 1\} = 2\psi(1) = 0.6826$; $P\{|Z| \leq 2\} = 2\psi(2) = 0.9544$; $P\{|Z| \leq 3\} = 2\psi(3) = 0.9974$;
$P\{|Z| \leq 1.96\} = 2\psi(1.96) = 0.9500$; $P\{|Z| \leq 2.58\} = 2\psi(2.58) = 0.9902$.

From above, we infer that area under the standard normal curve $N(0, 1)$ and bounded by the lines $z = \pm 1.96$ is 95%. This means that the area beyond the ordinates $z = \pm 1.96$ is 0.05 (i.e. 5%). Because of symmetry, there is 2.5% area in either tale.

We further infer that area under the standard normal curve $N(0, 1)$ and bounded by the ordinates $z = \pm 2.58$ is 99.02% which may be roughly treated as 99%. Thus limits $z = \pm 2.58$ exclude 1% area ($\frac{1}{2}$% in each tail) in $N(0, 1)$ curve.

Observations. Almost all z-values lie between $\mu \pm 3\sigma$ which are called **three-sigma** limits.

16-22. *One-tailed Area*

To find limits which set aside particular areas in one tail, we use inverse reading from the area-table of $N(0, 1)$. Thus, the probability 0.05 in the right tail corresponds to $z = 1.645$ and the probability 0.01 (area) in the right tail corresponds to $z = 2.326$.

16-23. *Some Standard Terms*

1. A Statistical Hypothesis (notated : H) is a statement about the nature of a universe or population. It is frequently stated in terms of a population parameter/s.

2. The Null Hypothesis (notated : H_0) is a statement about a population parameter which is to be tested. Thus H_0 is an H we need to test.

3. A Test Statistic (notated : TS) is a function of the data related to a sample. The value of TS, under H_0 written $(TS)_0$ helps reject or accept H_0. Generally, TS is a standardized variate.

4. Critical Regions. The set of points on the z-axis which constitute the tails may be denoted by C and is called *critical region* or *region of rejection.* Obviously, C = {Values of TS for which H_0 is rejected}

5. Level of Significance. The size (area) of the critical region is called a level of significance and is denoted by α. Obviously, α is the maximal probability of rejecting H_0, when H_0 is true. The most frequently used levels are $\alpha = 0.05$ (5%) and $\alpha = 0.01$ (1%). The critical regions at $\alpha = 0.05$ are

$$C = \{Z : |z| \geq 1.96\} \text{ for 2-tailed (both-sides) regions.}$$
$$C = \{Z : z \geq 1.645\}, \text{ for 1-tailed (right-side) region.}$$

The critical regions at $\alpha = 0.01$ (1%) are

$$C = \{Z : |z| \geq 2.58\} \text{ for 2-tailed (both-sides) regions.}$$
$$C = \{Z : z \geq 2.326\}, \text{ for 1-tailed (right-tailed) region.}$$

We summarize some of the frequently used critical values (See Fig. p. 486)

k :	1.96	1.65	1.28	0.126	0.062
$P\{|Z| \geq k\}$:	0.05	0.10	0.20	0.90	0.95

Convention. $\alpha = P\{Z \geq z_\alpha\}$: Area beyond (*i.e.* in front of) $Z = z_\alpha$ in $N(0, 1)$ curve.

6. The smallest significance level at which H_0 is rejected is called ***p*-value.**

16-24. Worked-out Numericals using N(0, 1)-Tabulations

Example 1. If $Y = \log_{10} X$ is $N(4, 2^2)$ and $\log_{10} 1202 = 3.08$, $\log_{10} 8318 = 3.92$, find $P\{1.202 < X < 83180000\}$.

Solution.
$$p = P\{1.202 < X < 8318 \times 10^4\} = P\{\log 1.202 < \log X < \log 8318 + 4\}$$
$$= P\{0.8 < Y < 7.92\} = P\{-1.96 < (Y-4)/2 < 1.96\}, \quad [Y \sim N(4, 4)]$$
$$= P\{-1.96 < Z < 1.96\} = 0.9500. \quad \text{[by Standardization } Z = (Y-4)/2]$$

Example 2. Let $X \sim N(\mu, \sigma^2)$, and $\Phi(k)$ be the c.d.f. of $N(0, 1)$.
(a) If $\sigma^2 = \mu^2 (\mu > 0)$, express $P\{X < -\mu \mid X < \mu\}$ in terms of $\Phi(k)$.
(b) If $\sigma^2 = h(\mu)$, $\mu > 0$, find the function h such that $P\{X \le 0)$ does not depend on μ.

Solution. (a) $p = P\{X < -\mu \mid X < \mu\} = P\{Z \le -2 \mid Z \le 0\}$, $\quad [Z = (X-\mu)/\sigma \text{ i.e. } X = \mu + \sigma Z]$
$$= P\{Z \le -2\} / P\{Z \le 0\} = 2\Phi(-2) \qquad [\because P(Z \le 0) = 1/2]$$

(b) Let $H(\mu) = P(X \le 0) = P(Z \le -\mu/\sqrt{h}) = \int_{-\infty}^{-\mu/\sqrt{h}} f(z)\, dz$ ...(1)

where $f(z) = (1/\sqrt{2\pi}) \exp(-\frac{1}{2}z^2)$ is independent of μ. Now $H(\mu)$ shall be independent of μ, if the upper limit in (1) is independent of μ, which requires $\mu/\sqrt{h}$ = const. = $1/k^{1/2}$, say. This gives $h = k\mu^2$. Thus, with the choice $h(\mu) = k\mu^2$. $P(X \le 0)$ does not depend on μ.

Example 3. Let $X \sim N(\mu, \sigma^2)$. Show that the maximum of $P\{c - h < X \le c + h\}$ occurs at $c = \mu$, $h > 0$.

Solution. We note that, $Z = (X - \mu)/\sigma$ is $N(0, 1)$, so that
$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}, f'(z) = \frac{-z}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \qquad \text{...(1)}$$
$$p(c) \equiv P\{c - h < X < c + h\} = P\left\{\frac{c-h-\mu}{\sigma} < \frac{X-\mu}{\sigma} < \frac{c+h-\mu}{\sigma}\right\}$$
$$= F_Z\left(\frac{c+h-\mu}{\sigma}\right) - F_Z\left(\frac{c-h-\mu}{\sigma}\right) \qquad [F_Z \text{ is c.d.f. of } Z]$$
$$p'(c) = \frac{1}{\sigma} f\left(\frac{c+h-\mu}{\sigma}\right) - \frac{1}{\sigma} f\left(\frac{c-h-\mu}{\sigma}\right) \qquad \text{...(2)}$$
$$p''(c) = \frac{1}{\sigma^2} f'\left(\frac{c+h-\mu}{\sigma}\right) - \frac{1}{\sigma^2} f'\left(\frac{c-h-\mu}{\sigma}\right)$$
$$= -\frac{(c+h-\mu)}{\sigma^3} f\left(\frac{c+h-\mu}{\sigma}\right) + \frac{(c-h-\mu)}{\sigma^3} f\left(\frac{c-h-\mu}{\sigma}\right), \quad \text{[by (1)]}$$
$$p'(c) = 0 \Rightarrow (c+h-\mu)^2 = (c-h-\mu)^2 \Rightarrow c = \mu$$
$$p''(\mu) = -\frac{h}{\sigma^3}\left\{f\left(\frac{h}{\sigma}\right) + f\left(-\frac{h}{\sigma}\right)\right\} < 0.$$

Hence at $c = \mu$, $p(c)$ is maximum.

Comments. This proves that $\{X = \mu\}$ is the *modal point* of $N(\mu, \sigma^2)$.

Problems with Solutions Provided at the End of the Text

1*. If $X \sim N(30, 5^2)$ and $Y \sim N(15, 10^2)$, show that

$$P\{26 \le X \le 40\} = P\{7 \le Y \le 35\}.$$

2*. Show that the probability that the number of heads in 2000 throws with a fair coin lies between 900 and 1000 is approximately $2F(2\sqrt{5}) - 1$, where $F(z)$ is the c.d.f. of $N(0, 1)$.

3*. If X is $N(30, 5^2)$, find the probability of
(a) $26 \le X \le 40$, (b) $|X - 30| > 5$, (c) $X \ge 42$, (d) $P(X \le 28)$.

4*. Student's final score is the value from $N(\mu, \sigma^2)$ and instructor's grading is as under : Grade A when score exceeds $\mu + \sigma$, grade B when score falls between μ and $\mu + \sigma$, grade C if score falls between $\mu - \sigma$ and μ, grade D if score falls between $\mu - 2\sigma$ and $\mu - \sigma$ and score F if score falls below $\mu - 2\sigma$. Calculate probability of each grade.

5*. If X is $N(50, 10^2)$ find $P\{Y \le 3137\}$ where $Y = X^2 + 1$, given that

$$\psi(0.6) = (\sqrt{2\pi})^{-1} \int_0^{0.6} (e^{-x^2/2} = 0.2258.$$

6*. If X is $N(1, 4)$ find the values of

$$p_1 = P\{|X - 2| < \tfrac{1}{2} | X > 0\},\ p_2 = P\{X < 0 \mid |X - 1| > \tfrac{1}{2}\}$$

7*. If X is $N(\mu, \sigma^2)$ for an interval of length $\tau > 0$, for what value of x is the probability $P\{x \le X \le x + \tau\}$ a maximum ?

Exercise 16(a)

1. Let X be a normal variate with p.d.f.
$$f(x) = 0.03989 \exp[-0.005\,x^2 + 0.5\,x - 12.5], \quad -\infty < x < \infty$$
Express $f(x)$ in the standard form and hence or otherwise find the mean and Var X.
[**Ans.** $\mu = 50, \sigma^2 = 100$]

2. Find the mean and S.D. of a p.d.f. $f(x) = Ce^{-(x^2 - 6x + 9)/24}$, $-\infty < x < \infty$ where C is constant.
[**Ans.** $\mu = 3, \sigma^2 = 12, C = (24\pi)^{-1/2}$]

3. Determine the constant C so that $f(x) = Ce^{-\frac{1}{3}x^2 + 2x}$, $-\infty < x < \infty$, satisfies the conditions of being a p.d.f. Obtain the values of μ and σ^2.
[**Ans.** $\mu = 8, \sigma^2 = 4, C = e^{-8}(8\pi)^{-1/2}$]

4. Show that the constant C can be selected so that $f(x) = Ce^{-(9x^2 - 12x + 13)}$, $-\infty < x < \infty$ is the p.d.f. of Normal variate. Find also the values of μ and σ^2. [**Ans.** $\mu = 2/4, \sigma^2 = 1/18, C = 3e^9/\sqrt{\pi}$.]

5. Show that a constant k can be selected so that $f(x) = k2^{-x^2}$, $-\infty < x < \infty$, satisfies the conditions of a normal p.d.f. Find also the values of proper μ and σ^2.
[**Ans.** $\mu = 0, \sigma^2 = (2 \ln 2)^{-1}, k = (\ln 2/\pi)^{1/2}$.

6. If X is $N(0, 1)$, show that for all $x > 0$

(a) $\Phi(-x) = 1 - \Phi(x)$; (b) $P\{-x \le X \le x\} = 2\Phi(x) - 1$

(c) $P\{|X - \mu| < k\} = 2\Phi(k/\sigma) - 1$; (d) $P\{|X| > x\} = 2[1 - \Phi(x)] = 2\Phi(-x)$.

7. (a) If X is $N(2, 1)$, find $P\{|X - 2| \le 1\}$.

(b) If X is $N(2, 2)$, express $P\{|X - 1| \le 2\}$ in terms of $N(0, 1)$ c.d.f.

[**Ans.** 0.6826, $F(2^{-1/2}) + F(3/\sqrt{2}) - 1$]

8. Let $X \sim N(0, 1)$, $Y \sim N(2, 4)$ and $Z \sim N(-1, 1)$ be independent. Compute the probability that exactly two of these three variates are less than zero. [**Ans.** 0.4328]

9. Let X be $N(12, 42)$. Find the probabilities

(i) $P(X \ge 20)$, (ii) $P(X \le 20)$ (iii) $P(X \le 12 | X \ge 0)$.

Find the constants a, b, c when $P(a < X < b) = 0.5$, $P(X < a) = 0.25$, $P(X > c) = 0.24$.

[**Ans.** 0.0228, 0.9772, 0.5; $a = 9.3$, $b = 14.7$, $c = 14.8$]

10. If X is $N(75, 25)$, find $P\{X > 80 | X > 77\}$. [**Ans.** 0.46]

11. If $X \sim N(10, ?)$ and $P(X > 12) = 0.1587$ show that $P\{9 \le X \le 11\} = 0.3830$, use $\Phi(1) = 0.8413$, $\Phi(-\frac{1}{2}) = 0.3085$.

12. If X is $N(11, 2.25)$, find the number k such that

(a) $P(X > k) = 0.3$ (b) $P(X > k) = 0.09$. [**Ans.** 11.79, 13.01]

13. If $X \sim N(5, 16)$ and $Y \sim N(4, 1)$ find k such that

(a) $P(X < k) = 0.919$ (b) $P\{|Y - 4| > k\} = 0.012$. [**Ans.** 10.6, 2.51]

14. If $\ell n\, X \sim N(1, 4)$ and $\ell n_e\, 2 = 0.693$, find $P\{1/2 < X < 2\}$. [**Ans.** 0.2427]

15. If X is $N(0, 1)$, use Chebyshev's inequality to estimate $P(|X| > 1.96)$. [**Ans.** 0.26 (crude)]

16. Let $X \sim N(\mu, \sigma^2)$.

(a) If $P(X < 89) = 0.90$, $P(X < 94) = 0.95$, find μ and σ^2.

(b) If $\sigma = 2$ and $P(X \le 3) = 0.8$, find μ.

(c) If $\mu = 0$ and $P(|X| > 8) = 0.3$, find σ.

(d) If $P\{(4X - 3) > 0\} = P\{(4 - 3X) = 0\} = 0.9$, find μ and σ.

(e) If $P\{|X - \mu| \ge c\} = P\{|X - \mu| \le c\}$, $c > 0$; find c.

(f) If $\sigma = 10$, $P\{X < 80.5\} = 0.3264$, find μ.

[**Ans.** (a) 72.46 and $(13.7)^2$, (b) 1.32, (c) 5.7, (d) 1.04, 0.227, (e) 0.675 σ, (f) 76]

17. (a) If $X \sim N(1, 1)$, find $P\{(X - X^2) > 0\}$. (b) If $X \sim N(\mu, \mu^2/9)$, find $P\{X < 0\}$.

[**Ans.** 0.3413, 0.0013]

18. If $X \sim N(5, 10)$ find $P(0.04) < (X - 5)^2 < 38.4\}$. [**Ans.** 0.90]

19. A normal distribution has mean 77. Find its S.D. If 20% of the area under the curve lies to the right of 90. [**Ans.** 15.5]

20. Show that the p.d.f. $f(x, u) = (2\pi u)^{-1/2} \exp(-x^2/2u)$, $-\infty < x < \infty$, of a $N(0, u)$ variate satisfies the heat equation $(\partial f/\partial u) = \frac{1}{2}(\partial^2 f/\partial x^2)$.

16-30. Some Standard Properties for $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}, -\infty < x < \infty$

1. *Mean (absolute) Deviation M.(a).D.* :

If M denotes the mean deviation of X about μ, then using definition

$$M = E(|X-\mu|) = \int_{-\infty}^{\infty} \frac{e^{-\frac{1}{2}(x-\mu)^2/\sigma^2}}{\sqrt{2\pi}\,\sigma} |x-\mu|\,dx = \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{\infty} |z| e^{-z^2/2}\,dz, \left[z = \frac{x-\mu}{\sigma}\right]$$

$$= \frac{2\sigma}{\sqrt{2\pi}} \int_0^{\infty} z e^{-\frac{1}{2}z^2}\,dz = \sigma\sqrt{\frac{2}{\pi}} \qquad \text{[Even integrand]}$$

Thus M.D. $= \sigma\sqrt{2/\pi} = (4/5)\sigma$.

2. *Mean deviation of order r* :

$$M_r = E(|X-\mu|^r) = \int_{-\infty}^{\infty} f(x)|x-\mu|^r\,dx \qquad \left[\text{Put } \frac{x-\mu}{\sigma} = z\right]$$

$$= \frac{\sigma^r}{\sqrt{2\pi}} \int_{-\infty}^{\infty} |z|^r e^{-\frac{1}{2}z^2}\,dz \qquad \text{(Integrand is even)}$$

$$= \frac{2\sigma^r}{\sqrt{2\pi}} \int_0^{\infty} z^r . e^{-\frac{1}{2}z^2}\,dz = \frac{\sigma^r (2)^{r/2}}{\sqrt{\pi}} \int_0^{\infty} e^{-t} (t)^{\frac{r+1}{2}-1}\,dt$$

$$= \frac{\sigma^r (2)^{r/2}}{\sqrt{\pi}} \Gamma\left(\frac{r+1}{2}\right).$$

Note. $$M_{2r} \equiv \mu_{2r} = \frac{\sigma^{2r}\, 2^r}{\sqrt{\pi}} \Gamma\left(\frac{2r+1}{2}\right).$$

3. *Quartile Deviation* :

Let Q_1, Q_2 $(= \mu)$ and Q_3 be the lower, middle and upper quartiles for $N(\mu, \sigma^2)$. By definition

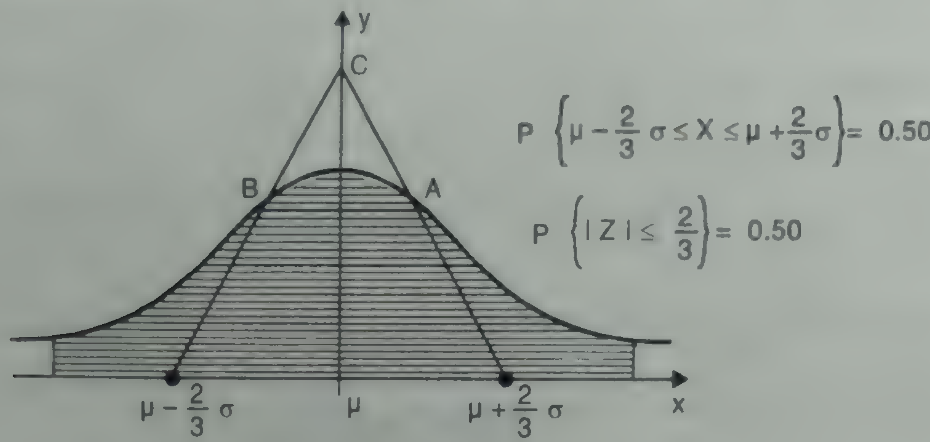
$$\int_{\mu}^{Q_3} f(x)\,dx = 0.25 \Rightarrow \int_0^q \frac{e^{-z^2/2}}{\sqrt{2\pi}}\,dz = 0.25, \qquad \text{where } z = \frac{x-\mu}{\sigma}, q = \frac{Q_3-\mu}{\sigma}.$$

From the Table of Normal areas this gives, by inverse reading, $q = 0.674 \Rightarrow Q_3 - \mu = \sigma(0.674)$. Symmetrically, $\mu - Q_1 = \sigma(0.674)$ so that $Q = \frac{1}{2}(Q_3 - Q_1) = (0.674)\,\sigma \doteq (2/3)\sigma$. [Quartile deviation]

Probable error. It is an old name for quartile deviation Q. This may also be defined by $P\{\mu - Q < X < \mu + Q\} = 1/2$. Now

$$1/2 = P(\mu - Q < X < \mu + Q) = P(-Q/\sigma) < Z < (Q/\sigma)\} = 2\psi(Q/\sigma)$$

From tables, $\psi(Q/\sigma) = 1/4 \Rightarrow Q/\sigma = 0.674$ so $Q \doteq (2/3)\sigma$.



4. *Median :*

Let $X = m$ be the medonal value ; then by definition

$$\frac{1}{2} = \int_m^\infty f(x)\,dx = \int_m^\mu f(x)\,dx + \int_\mu^\infty f(x)\,dx = \int_m^\mu f(x)\,dx + \int_0^\infty \frac{e^{-z^2/2}}{\sqrt{2\pi}}\,dz = \int_m^\mu f(x)\,dx + \frac{1}{2} \quad \text{...(i)}$$

where $z = (x - \mu)/\sigma$ and $\int_0^\infty f(z)\,dz = 1/2$. So,

$$\frac{1}{2} = \int_m^\mu f(x)\,dx + \frac{1}{2} \Rightarrow \int_m^\mu f(x)\,dx = 0 \Rightarrow m = \mu \qquad [\because f(x) > 0].$$

Aliter. Since $N(\mu, \sigma^2)$ is symmetric about $X = \mu$, it follows that μ is the medonal value. In fact $P(X < \mu) = 1/2 = P(X > \mu)$:

$$P(X > \mu) = \int_\mu^\infty \frac{\exp[-(x-\mu)^2/2\sigma^2]}{\sigma\sqrt{2\pi}}\,dx = \int_0^\infty \frac{e^{-z^2/2}}{\sqrt{2\pi}}\,dz = \frac{1}{2} = P(X < \mu).$$

5. *Mode :*

The modal value is the variate-value that corresponds to the maximum probability. Here

$$y = f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2};\ \ell n\, y = \ell n\,(\sigma\sqrt{2\pi})^{-1} - (x-\mu)^2/2\sigma^2.$$

Differentiating w.r.t. 'x' we obtain

$$(y'/y) = -(x-\mu)/\sigma^2;\quad (y''/y) - (y'/y)^2 = -(1/\sigma^2). \quad \text{...(1)}$$

Putting $y' = 0$ we get $x = \mu$ and from Eq. 1(b) $y'' = -(y/\sigma^2) < 0$, since $y = (\sigma\sqrt{2\pi})^{-1}$ at $x = \mu$. Thus $X = \mu$ is the mode of $N(\mu, \sigma^2)$.

Maximum Probability : max $f(x) = (\sigma\sqrt{2\pi})^{-1}$, [at $x = \mu$, i.e. at mode].

Observation : Mean = Mode = Median (each = μ).

6. *Points of Inflexion :*

Analytically, *inflexion points* are defined by the solutions to equation $(d^2y/dx^2) = 0$, $(dy/dx \neq 0)$. Here

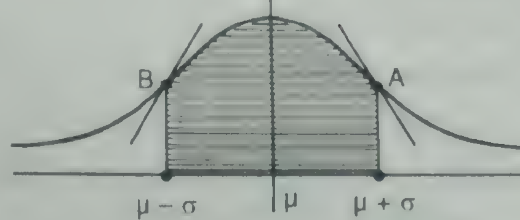
$$y = f(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-(x-\mu)^2/2\sigma^2}, \text{ so } \ \ell n\, y = \ell n\,(\sigma\sqrt{2\pi})^{-1} - (x-\mu)^2/2\sigma^2 \quad \text{...(1)}$$

Differentiating twice w.r.t. x we obtain

$y'/y = -(x-\mu)/\sigma^2$; $(y''/y)-(y'-y)^2 = -1/\sigma^2$ ...(2)

Putting $y''=0$, we get from (2), $(x-\mu)^2 = \sigma^2 \Rightarrow x = \mu \pm \sigma$. ...(3)

Putting for x from (3) into (1) we get, $y = (\sqrt{2\pi e}\,\sigma)^{-1}$.



Thus the points of inflexion are given by $A[\mu+\sigma, (\sigma\sqrt{2\pi e})^{-1}]$ and $B[\mu-\sigma, (\sigma\sqrt{2\pi e})^{-1}]$.

$$p = P\{\mu-\sigma \le X \le \mu+\sigma\} = P(-1 \le Z \le 1) = 0.6806.$$

7. *Angle between Flex-tangents :*

Let m, m' be the slopes of the tangents at inflexions points A and B; then from §5(2)

$$m = -(\sigma^2\sqrt{2\pi e})^{-1}, \quad m' = (\sigma^2\sqrt{2\pi e})^{-1} = -m$$

$$\therefore \qquad \tan\theta = \frac{m'-m}{1+m'm} = \frac{2/\sigma^2\sqrt{2\pi e}}{1-(1/\sigma^4 . 2\pi e)} = \frac{2\sigma^2\sqrt{2\pi e}}{2\pi e\sigma^4 - 1}$$

The flex-tangents at A and B are

$$y-(1/\sigma\sqrt{2\pi e}) = m(x-\mu-\sigma); \; y-(1/\sigma\sqrt{2\pi e}) = m'(x-\mu+\sigma). \qquad ...(1)$$

Subtracting 1(b) from 1(a) gives $x = \mu$ and putting $x = \mu$ into 1(a) gives $y = 2/\sigma\sqrt{2\pi e}$.

Hence the point of intersection of the flex tangents is $C(\mu, 2/\sigma\sqrt{2\pi e})$.

16-31. *Worked-out Problems*

***Example* 1.** If Z is $N(0, 1)$, then

$$P\{|Z| \ge t\} \le \sqrt{\frac{2}{\pi}}\frac{e^{-t^2/2}}{t}, \; \forall\, t > 0.$$

Solution.

$$P\{Z \ge t\} = \frac{1}{\sqrt{2\pi}}\int_t^\infty e^{-x^2/2}\,dx \qquad \left[x > t \Rightarrow \frac{t}{x} < 1\right]$$

$$\le \frac{1}{\sqrt{2\pi}}\int_t^\infty \frac{x}{t} e^{-x^2/2}\,dx$$

$$= \frac{1}{\sqrt{2\pi}}\int_{t^2/2}^\infty \frac{1}{t} e^{-u}\,du \qquad \left[\frac{1}{2}x^2 = u\right]$$

$$= \frac{1}{\sqrt{2\pi}}\frac{1}{t}e^{-t^2/2} \qquad ...(1)$$

The symmetry of $N(0, 1)$ yields

$$P\{|Z| \ge t\} = 2P\{Z \ge t\} \le \sqrt{\frac{2}{\pi}}\frac{e^{-t^2/2}}{t}, \text{ by (1).}$$

***Example* 2.** Show that for normal distribution, the quartile deviation, the mean deviation and the standard deviation are approximately in the ratio 10 : 12 : 15.

Solution. We know that, for $N(\mu, \sigma^2)$; Q.D. = $(2/3)\sigma$, M.D. = $(4/5)\sigma$, S.D. = σ

$\therefore$ Q.D. : M.D. : S.D. = $(2/3)\sigma : (4/5)\sigma : \sigma = 10 : 12 : 15$.

Note. For complete solution; these results have to be derived instead of being quoted.

***Example* 3.** Let $W = Z^2$, where Z is $N(0, 1)$, show that $F_W(w) = 2\Phi(w^{1/2}) - 1$.

Hence deduce the median and the interquartile range for W from that of $N(0, 1)$ tables.

Solution. $F_W(w) = P\{W \le w\} = P(Z^2 \le w\} = P\{-\sqrt{w} \le Z \le \sqrt{w}\} = 2\psi(\sqrt{w}) = 2\Phi(w^{1/2}) - 1.$ ...(1)

Since $\psi(k) + \frac{1}{2} = \Phi(k).$ [Draw Figure]

Deductions. If $W = m$ is the median, then $F(m) = \frac{1}{2}$, and Eq. (1) gives $\Phi(\sqrt{m}) = \frac{1}{2}(1 + \frac{1}{2}) = 3/4 = 0.75$. Scanning the area table $N(0, 1)$, we find $\sqrt{m} = 0.675$, so that $m = 0.456$.

If Q_1 and Q_2 are the lower and upper quartiles of W. Using $F(Q_1) = \frac{1}{4}, F(Q_3) = 3/4$, Eq. (1) provides $\Phi(\sqrt{Q_3}) = 7/8 = 0.875, \Phi(\sqrt{Q_1}) = 5/8 = 0.625$.

Scanning once again $N(0, 1)$ areas, we find $\sqrt{Q_3} = 1.15$, $\sqrt{Q_1} = 0.32$,

$$\therefore \quad Q_3 - Q_1 = (1.15)^2 - (0.32)^2 = 1.3225 - 0.1024 = 1.2201.$$

Note. Later on we shall see that $W = [N(0, 1)]^2$ is $\chi^2_{(1)}$. So we can read the values of median and quartiles immediately from χ^2-tables.

Problems with Solutions Provided at the End of the Text

1*. If two Normal universes A and B have the same total frequency but S.D. of universe A is k times that of universe B, show that the maximum frequency of universe A is $(1/k)$ times that of universe B.

2*. 5000 variates are normally distributed with mean 50 and probable error 13.49. Without using tables, find Q_1, Q_2, $\hat{x}$ (mode), M.D., S.D. Find also the variate value for which the cumulative frequency is 1250.

16-40. *Moment Generating Function. Ch. Function & Cumulants*

$$M(t : X) = e^{\mu t + (\sigma^2 t^2/2)}; \qquad M(t : Z) = e^{t^2/2} \qquad [\S 8.16(8)]$$

$$M(it) = e^{i\mu t - (\sigma^2 t^2/2)}; \qquad M(it : Z) = e^{-t^2/2} \qquad [\S 9\text{-}20]$$

$$K(t : X) = \ln M(t : X) = \mu t + \tfrac{1}{2}\sigma^2 t^2 \Rightarrow \Sigma k_r(t^r/r!) = \mu t + (\sigma^2 t^2/2).$$

So comparing the Coeffts. of $t^r/r!$ on either side we get $k_1 = \mu$, $k_2 = \sigma^2$, $k_r = 0$ for $r = 3, 4, \ldots$

16-41. Worked-out Problems

Example **1.** If X is $N(0, 1)$ and Y is $N(0, 1)$, prove that

(a) Var $(\sin X) >$ Var $(\cos X)$

(b) $\mathsf{E}|X-Y| \le \sqrt{8/\pi}$, for any relation between X and Y.

Solution. (a) Here $\mathsf{E}(e^{itX}) = e^{-\frac{1}{2}t^2} \Rightarrow \mathsf{E}(\cos X + i \sin X) = e^{1/2}$; $\mathsf{E}(\cos 2X + i \sin 2X) = e^{-2}$.

Thus,

$\mathsf{E}(\cos X) = e^{-1/2}, \mathsf{E}(\sin X) = 0, \mathsf{E}(\cos 2X) = e^{-2}, \mathsf{E}(\sin 2X) = 0$ (Equating real and imaginary parts)

Var $(\sin X) = \mathsf{E}(\sin^2 X) - [\mathsf{E}(\sin X)]^2 = \mathsf{E}[(1-\cos 2X)/2] = (1-e^{-2})/2 = 0.4323.$

Var $(\cos X) = \mathsf{E}(\cos^2 X) - [\mathsf{E}(\cos X)]^2 = \mathsf{E}[(1+\cos 2X)/2] - e^{-1} = \frac{1}{2}(1-e^{-1})^2 = 0.1997.$

Thus, Var $(\sin X) >$ Var $(\cos X)$.

(ii)
$$\mathsf{E}\,|X| = \int_{-\infty}^{\infty} |x| \frac{e^{-\frac{1}{2}x^2}}{\sqrt{2\pi}}\,dx = \frac{2}{\sqrt{2\pi}} \int_0^{\infty} x e^{-\frac{1}{2}x^2}\,dx = \sqrt{\frac{2}{\pi}}$$

$$|X-Y| \le |X| + |Y| \Rightarrow \mathsf{E}|X-Y| \le \mathsf{E}|X| + \mathsf{E}|Y| = 2.\sqrt{2/\pi} = \sqrt{8/\pi}.$$

***Example* 2.** Let $Z = XY$, where X and Y are independent variates with $P(X = \pm 1) = \frac{1}{2}$ and $Y \sim N(0, 1)$. Show that Z is $N(0, 1)$, that Y and Z are uncorrelated though *not* independent.

Solution.
$$M(t:Z) = \mathsf{E}(e^{tZ}) = \mathsf{E}(e^{tXY}) = \mathsf{E}[\mathsf{E}(e^{tXY} \mid X)] = \mathsf{E}[e^{\frac{1}{2}X^2t^2}] \quad \text{[By double-E Rule]}$$
$$= \tfrac{1}{2}e^{\frac{1}{2}t^2(1)^2} + \tfrac{1}{2}e^{\frac{1}{2}t^2(1)^2} = e^{t^2/2}.$$

Thus $Z \sim N(0, 1)$

Now $\mathsf{E}(X) = 1.(\frac{1}{2}) + (-1)(\frac{1}{2}) = 0$; $\mathsf{E}(YZ) = \mathsf{E}(XY^2) = \mathsf{E}(X)\mathsf{E}(Y^2) = 0.$

$$\text{Cov}(Y, Z) = \mathsf{E}(YZ) - \mathsf{E}(Y)\mathsf{E}(Z) = 0 \Rightarrow Y \text{ and } Z \text{ are uncorrelated.}$$

Since Y and Z are $N(0, 1)$; the *point* probability $P(Y + Z = 0) = 0$.

Now assuming Y and Z are indep., we have

$0 = P[Y+Z=0] = P[Y(1+X)=0] = P[(Y=0) \uplus (X=-1)] = P(Y=0) + P(X=-1) = 0 + \frac{1}{2}$

The obvious contradiction implies that Y and Z are *not* independent.

***Example* 3.** For $|\mu| < \infty$, $0 < p < 1$, let X be a mixed variate such that $f(x) = p(\sigma\sqrt{2\pi})^{-1} e^{-(x-\mu)^2/2\sigma^2}$, $-\infty < x < \infty$; omitting non-neg integers ; $f(x) = 0$, otherwise

$$g(x) = (1-p)e^{-\lambda}\lambda^x / x!, \quad x = 0, 1, 2, 3, \ldots$$

Find the m.g.f. and hence or otherwise obtain Var (X).

Solution. Using the m.g.f. of $N(\mu, \sigma^2)$ and Pois (λ) we get

$$\mathsf{E}(e^{tX}) = p\int_{-\infty}^{\infty} e^{tx} \frac{\exp[-(x-\mu)^2/2\sigma^2]}{\sigma\sqrt{2\pi}}\,dx + (1-p)\sum_{x=0}^{\infty} e^{tx}\frac{e^{-\lambda}\lambda^x}{x!}$$

$\therefore \quad M(t) = p[\text{m.g.f. of } N(\mu, \sigma^2)] + (1-p)[\text{m.g.f. of Pois.}(\lambda)] = pe^{\mu t + \frac{1}{2}\sigma^2 t^2} + (1-p)\, e^{\lambda(e^t - 1)}$.

Expanding the exponentials involved, writing $q = 1 - p$, we get

$$M(t) = p[1 + (\mu t + \tfrac{1}{2}\sigma^2 t^2) + \tfrac{1}{2}(\mu t + \tfrac{1}{2}\sigma^2 t^2)^2 + \ldots] + qe^{\lambda(t + t^2/2 + \ldots)}$$

$$= p[1 + \mu t + (\sigma^2 + \mu^2) t^2/2 + \ldots] + q\,[1 + \lambda\,(t + t^2/2 + \ldots) + \tfrac{1}{2}\lambda^2 (t + t^2/2 + \ldots)^2 + \ldots]$$

$$\mathsf{E}(X) = \text{Coeff. of } t = p\mu + q\lambda\,; \mathsf{E}(X^2) = \text{Coeff. of } \tfrac{1}{2}t^2 = p(\sigma^2 + \mu^2) + q(\lambda + \lambda^2)$$

$$\therefore \quad \text{Var}(X) = p(\sigma^2 + \mu^2) + q(\lambda + \lambda^2) - (p\mu + q\lambda)^2 = p\sigma^2 + q\lambda + pq(\mu - \lambda)^2.$$

Problems with Solutions Provided at the End of the Text

1*. If X is $N(0, 1)$, find the p.d.f. of $Y = k\,X^2$ and $Z = 2Y$.

2*. If $M_X(t) = e^{3t + 8t^2}$, find $P(-1 < X < 9)$.

3*. Let X be $N(\mu, 1)$ and $Y = [1 - \Phi(x)]/\varphi(x)$, where Φ and φ denote the c.d.f. and p.d.f. of $N(0, 1)$ variates respectively. Prove that $\mathsf{E}(Y) = \mu^{-1}$.

4*. If X, Y, Z are i.i.d, $N(0, 1)$ variates, Show that

$$W = (X + YZ)\big/\sqrt{1 + Z^2} \sim N(0, 1).$$

5*. Let X be $N(\mu, \sigma^2)$. Define (i.e. generate or simulate) a r.v. Y that has the same distribution as X and is negatively correlated with X.

6*. A random variable X has $\mathsf{E}(X^{2r}) = 2r!/2^r\,(r\,!)$, $\mathsf{E}(X^{2r-1}) = 0$, $r = 1, 2, 3, \ldots$ Find the m.g.f. and p.d.f. of X.

7*. Prove that if the independent variates X and Y have the p.d.fs. $(h/\sqrt{\pi})\,e^{-h^2 x^2}$ and $(k/\sqrt{\pi})\,e^{-k^2 y^2}$, then the variate $U = X + Y$ has the density $(l/\sqrt{\pi})\,e^{-l^2 u^2}$ where $l^{-2} = h^{-2}\,k^{-2}$.

8*. Let X be $N(0, 1)$. Find the characteristic function of $(X - a^2)$, where a is constant. Deduce : $k_r = 2^{r-1}(1 + ra^2)\,.\,(r-1)\,!$ where k_r is the rth cumulant of $(X - a)^2$.

16·50. *Moments of Normal Distribution*

1. ***Direct Evaluation of Central Moments.*** For $X \sim N(\mu, \sigma^2)$, definition of μ_r gives

$$\mu_r = \mathsf{E}(X - \mu)^r = \int_{-\infty}^{\infty} \frac{e^{-(x-\mu)^2/2\sigma^2}}{\sigma\sqrt{2\pi}}\,(x - \mu)^r\,dx = \int_{-\infty}^{\infty} \frac{e^{-\frac{1}{2}z^2}}{\sqrt{2\pi}}\,\sigma^r z^r\,dz. \qquad \left[z = \frac{x - \mu}{\sigma}\right]$$

When $r = 2n + 1$ (odd integer) the integrand is an odd function of z and hence the integral vanishes. This gives $\mu_{2n+1} = 0$. When $r = 2n$ (even integer), the integrand is an even function of z and the above integral can be expressed as

$$\mu_{2n} = 2\sigma^{2n} \int_0^{\infty} \frac{e^{-\frac{1}{2}z^2}}{\sqrt{2\pi}}\,z^{2n}\,dz = \frac{2^n\,\sigma^{2n}}{\sqrt{\pi}} \int_0^{\infty} e^{-u} u^{(n+1/2)-1}\,du, \qquad [u = \tfrac{1}{2}z^2]$$

$$= (2^n\,\sigma^{2n}/\sqrt{\pi})\,\Gamma(n + \tfrac{1}{2}).$$

$$\therefore \quad \mu_{2n} = \frac{(2n)!}{n!}\left(\frac{\sigma^2}{2}\right)^n, \quad \left[\Gamma\left(n+\frac{1}{2}\right) = \frac{(2n)!\sqrt{\pi}}{(n!)(2)^{2n}}\right] \quad ...(1)$$

Further, $$\frac{\mu_{2n}}{\mu_{2n-2}} = \frac{(2n).(2n-1)}{n}\left(\frac{\sigma^2}{2}\right) \Rightarrow \mu_{2n} = (2n-1)\sigma^2 . \mu_{2n-2}$$

Thus, $\mu_{k+2} = (k+1)\sigma^2 \mu_k$ (Moment Recurrence Formula), $(k \equiv 2n-2)$ ...(2)

2. *Moments through M.G.F.*

The m.g.f. of centred variate $N(0, \sigma^2)$

$$M(t : X - \mu) = e^{-\mu t} M(t : X) = e^{\sigma^2 t^2/2} \qquad [\S\ 8\text{-}16(6)]$$

Using power series for either side we get

$$\sum_{r=0}^{\infty} \mu_r \frac{t^r}{r!} = \sum_{r=0}^{\infty} \frac{(\frac{1}{2}\sigma^2 t^2)^r}{r!} = \sum_{r=0}^{\infty} \frac{\sigma^{2r}(2r)!}{r!(2)^r} \cdot \frac{t^{2r}}{(2r)!} \quad ...(2)$$

There are no odd-powers of t in (2) and hence $\mu_{2r+1} = 0$. Now equate the Coeff. of $t^{2r}/(2r)$! in (2) to get

$$\mu_{2r} = \frac{(2r)!}{r!}\left(\frac{\sigma^2}{2}\right)^r.$$

In particular, $\mu_2 = \sigma^2$; $\mu_4 = 3\sigma^4$. ...(3)

3. *Factorial moments.* Using relations between μ_r and $\mu_{(r)}$ we get

$$\mu_{(2)} = \mu_2 = \sigma^2; \quad \mu_{(3)} = \mu_3 - 3\mu_2 = -3\sigma^2 ; \ \mu_{(4)} = \mu_4 - 6\mu_3 + 11\mu_2 = 3\sigma^4 + 11\sigma^2 = \sigma^2(11 + 3\sigma^2).$$

16-51. *Romanvosky Moment Recurrence Formula*

$$\mu_{2n+2} = \sigma^2 \mu_{2n} + \sigma^3 [d\mu_{2n} / d\sigma] \quad ...(1)$$

Proof. $\mu_{2n} = \int_{-\infty}^{\infty} (x-\mu)^{2n} f(x)\,dx$, $f(x) = (\sigma\sqrt{2\pi})^{-1} e^{-(x-\mu)^2/2\sigma^2}$. [Definition]

We assume that this integral is uniformly convergent so that differentiation under integral sign is permissible. Thus

$$\frac{d}{d\sigma}(\mu_{2n}) = \int_{-\infty}^{\infty} (x-\mu)^{2n} \frac{d}{d\sigma} f(x)\,dx$$

As $$\ln f(x) = c - \ln \sigma - \frac{(x-\mu)^2}{2\sigma^2}; \quad \frac{d}{d\sigma} f(x) = \frac{1}{\sigma^3}[(x-\mu)^2 - \sigma^2] f(x).$$

$$\therefore \quad \frac{d}{d\sigma}\mu_{2n} = \frac{1}{\sigma^3}\int_{-\infty}^{\infty} f(x)[(x-\mu)^{2n+2} - \sigma^2 (x-\mu)^{2n}]\,dx = \frac{1}{\sigma^3}[\mu_{2n+2} - \sigma^2 \mu_{2n}].$$

Thus, by transfer of terms, the result (1) follows.

Another Proof. Since $\mu_{2n} = [2^n \Gamma(n+\frac{1}{2})/\sqrt{\pi}] . \sigma^{2n}$, [§ 16-50 (1)]

$$\therefore D(\mu_{2n}) = [2^n \Gamma(n+\tfrac{1}{2})/\sqrt{\pi}](2n)\sigma^{2n-1} = (2n/\sigma)\mu_{2n}. \qquad [D = d/d\sigma] \quad ...(i)$$

We now substitute (i) into R.H.S. of (1) and use moment recurrence

$$\sigma^2\mu_{2n} + \sigma^3(d\mu_2/d\sigma) = \sigma^2\mu_{2n} + 2n\sigma^2\mu_{2n} = (2n+1)\sigma^2\mu_{2n} = \mu_{2n+2}. \qquad \text{[by §16-50(2)]}$$

Pearson's Coefficients

Since $D(\mu_2) = 2\sigma$, Eq. (1) gives

$$\mu_4 = \sigma^4 + \sigma^3(2\sigma) = 3\sigma^4 \Rightarrow \beta_1 = 0 \text{ and } \beta_2 = \mu_4/\mu_2^2 = 3$$

Example : Folded Normal distribution. If $X \sim N(0, \sigma^2)$, find the density of folded normal r.v. $Y = |X|$ and evaluate β_2.

Solution. $F_Y(y) = P\{|X| \le y\} = P\{-y \le X \le y\} = 2P\{0 \le X \le y\} = 2\int_0^y f_X(x)\,dx$

By DUIS : $F_Y(y) = 2f_X(y) = 2(\sigma\sqrt{2\pi})^{-1}\exp(-y^2/2\sigma^2), 0 < y < \infty.$ (1)

$$\mathrm{E}(Y^r) = \frac{2}{\sigma\sqrt{2\pi}}\int_0^\infty y^r e^{-y^2/2\sigma^2}dy = \frac{\sigma^r(\sqrt{2})^r}{\sqrt{\pi}}\int_0^\infty e^{-z} z^{(r+1)/2-1}dz \qquad \left[z = \frac{y^2}{2\sigma^2}\right]$$

$$= [\sigma^r(\sqrt{2})/\sqrt{\pi}]\,\Gamma[(r+1)/2].$$

$$\therefore \quad \mathrm{E}(Y) = \sigma(2/\pi)^{1/2}, \mathrm{E}(Y^2) = \sigma^2,\ \mathrm{E}(Y^3) = (\sqrt{8}/\pi)\sigma^3,\ \mathrm{E}(Y^4) = 3\sigma^4.$$

$$\mathrm{Var}(Y) = \mathrm{E}(Y^2) - \mathrm{E}^2(Y) = \sigma^2\left(1-\frac{2}{\pi}\right); \mu_3(y) = \mu_3' - 3\mu_2'.\mu_1' - 2\mu_1'^3 = \sqrt{\frac{2}{\pi}}\left(\frac{4}{\pi}-1\right)\sigma^3.$$

$$\mu_4(y) = \mu_4' - 4\mu_3'.\mu_1' + 6\mu_2'.\mu_1'^2 - 3(\mu_1')^4 = \sigma^4[3-(4/\pi)-(12/\pi)^2].$$

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{2}{\pi}\left(\frac{4}{\pi}-1\right)^2\Big/\left(1-\frac{2}{\pi}\right)^3; \ \beta_2 = \left(3-\frac{4}{\pi}-\frac{12}{\pi^2}\right)\left(1-\frac{2}{\pi}\right)^{-2}$$

$$\gamma_1 = \sqrt{\beta_1} = \left(\frac{4}{\pi}-1\right)\left(1-\frac{2}{\pi}\right)^{-1}\left(\frac{\pi}{2}-1\right)^{-1/2} = 0.9953,\ \beta_2 = \left(3-\frac{4}{\pi}-\frac{12}{\pi^2}\right)\left(1-\frac{2}{\pi}\right)^{-2} = 3.8692$$

Problems with Solutions Provided at the End of the Text

1*. For a certain $N(\mu, \sigma^2)$, the first moment about 10 is 40 and the 4th moment about 50 is 48. What is the mean and S.D. of $N(\mu, \sigma^2)$.

2*. If $X \sim N(\mu, \sigma^2)$, find the mean and variance of $Y = \frac{1}{2}[(X-\mu)/\sigma]^2$.

16.52. *Linear Combinations of Normal Variates*

If $X \sim N(\mu, \sigma^2)$ and $Y \sim N(\mu', \sigma'^2)$ are *independent* variates then $aX + bY + c$ is $N(a\mu + b\mu' + c, a^2\sigma^2 + b^2\sigma'^2)$, where a, b, c are constant. [Proof in § 8-18 (1)]

Cor. Reproductive Property or Addition Theorem : $X + Y \sim X(\mu + \mu', \sigma^2 + \sigma'^2)$.

Note. Extension to n independent variates is obvious.

16-53. Distribution of Mean of n i.i.d. Normal Variates

Let $X_i \sim N(\mu, \sigma^2)$, $i = 1, 2, \ldots, n$ be i.i.d. variates and $\overline{X} = \Sigma X_i / n$.

$$M(t : \overline{X}) = M(t : \Sigma X_i / n) = M(t/n : X_1 + X_2 + \ldots + X_n) = [M(t/n : X_i)]^n \quad (\because X_i \text{ are i.i.d.})$$

$$= [e^{\mu(t/n)} + \tfrac{1}{2}\sigma^2(t/n)^2]^n = e^{\mu t + \frac{1}{2}\sigma^2(t^2/n)}$$

This shows that $\overline{X} \sim N(\mu, \sigma^2/n)$.

16-54. Independence of Sample Mean & Sample Variance of N(μ, σ²) Population

Let $X_1, X_2, \ldots X_n$ be i.i.d. $N(0, 1)$ variates. Then mean and variance are defined by

$$\overline{X} = \left(\sum_{i=1}^{n} X_i\right)\Big/ n, \quad \hat{S}^2 = \left[\sum_{i=1}^{n}(X_i - \overline{X})^2\right]\Big/(n-1).$$

We show that $\overline{X}$ is independent of $(X_i - \overline{X})$. We use MGF technique.

$$M(t_1, t_2) = E\{e^{t_1 \overline{X} + t_2(X_i - \overline{X})}\} = E\{e^{(t_1 - t_2)\overline{X} + t_2 X_i}\} \qquad \text{[Def. of 2-dim m.g.f.]}$$

$$= E\left\{\exp\left[\left(\frac{t_1 - t_2}{n}\right)\Sigma X_j + \left(t_2 + \frac{t_1 - t_2}{n}\right)X_i\right]\right\}, \quad [\Sigma X_j \text{ consists of } (n-1) \text{ terms, has no } X_i\text{-term}]$$

$$= E\{\exp[(\theta Y) + (t_2 + \theta)X_i], \ [\theta = (t_1 - t_2)/n, \ Y = \Sigma X_j]$$

$$= E\{e^{\theta Y}\}.E\{e^{(t_2 + \theta)X_i}\}, \quad [\text{by independence of } X_i \text{ and } Y]$$

Now $Y \sim N[(n-1)\mu, (n-1)\sigma^2]$, $X_i \sim N(\mu, \sigma^2)$, we use m.g.f. of normal variate to obtain

$$M(t_1, t_2) = \exp[(n-1)\mu\theta + \tfrac{1}{2}(n-1)\sigma^2\theta^2].\exp[\mu(t_2 + \theta) + \tfrac{1}{2}\sigma^2(t_2 + \theta)^2]$$

$$= \exp\{\mu t_1 + \tfrac{1}{2}\sigma^2[(n-1)\theta^2 + (t_2^2 + 2t_2\theta + \theta^2)]\}$$

$$= \exp[\mu t_1 + \tfrac{1}{2}(\sigma^2/n)t_1^2].\exp\{0t_2 + \tfrac{1}{2}[(n-1)/n]\sigma^2 t_2^2\}$$

$$= M(t_1 : \overline{X}).M(t_2 : X_i - \overline{X}), \qquad [E(X_i - \overline{X}) = 0, \quad \text{Var}(X_i - \overline{X}) = \tfrac{1}{2}[(n-1)\sigma^2/n]$$

By Factorization Theorem, we conclude that $\overline{X}$ is independent of $X_i - \overline{X}$, and these are $N(\mu, \sigma^2/n)$ and $N[0, (n-1)\sigma^2/n]$ variates.

As suffix is arbitrary, $\overline{X}$ is independent of each $X_1 - \overline{X}, X_2 - \overline{X}, \ldots, X_n - \overline{X}$ and hence $\overline{X}$ is independent of the function $\hat{S}^2 = [\Sigma(X_i - \overline{X})^2]/(n-1)$.

16.55. Mean and Variance of a Truncated Normal Distribution

The density of $N(\mu, \sigma^2)$ is truncated for $X \le a$ and for $X \ge b$ and we evaluate mean and variance of the truncated distribution in $a \le X \le b$.

Put $z = (x - \mu)/\sigma$ so that $x = \mu + \sigma z$. Also $a_1 = (a - \mu)/\sigma$, $b_1 = (b - \mu)/\sigma$. Now

$$P\{a \le X \le b\} = P = \left\{\frac{a-\mu}{\sigma} \le \frac{X-\mu}{\sigma} \le \frac{b-\mu}{\sigma}\right\} = P\{a_1 \le Z \le b_1\} = \int_{a_1}^{b_1} \varphi(z)\,dz = \Phi(b_1) - \Phi(a_1) = 1/\lambda \text{ say}$$

The p.d.f. of truncated distribution is

$$g(x) = f(x)/P(a \le X \le b) = \lambda f(x), \qquad \varphi(x) = e^{-x^2/2}(\sqrt{2\pi})^{-1}.$$

$$\therefore \; \mathsf{E}(X^k) = \int_a^b g(x)\, x^k dx = \lambda \int_{a_1}^{b_1} (\mu + \sigma z)^k \varphi(z)\, dz \qquad \ldots(1)$$

$$\mathsf{E}(X) = \lambda \int_{a_1}^{b_1} (\mu + \sigma z)\varphi(z)\, dz = \lambda\mu \int_{a_1}^{b_1} \varphi(z)\, dz + \lambda\sigma \int_{a_1}^{b_1} z\varphi(z)\, dz$$

$$= \mu + \lambda\sigma \int_{a_1}^{b_1} \frac{ze^{-z^2/2}}{\sqrt{2\pi}}\, dz = \mu - \lambda\sigma\left(\frac{e^{b_1^2/2}}{\sqrt{2\pi}} - \frac{e^{a_1^2/2}}{\sqrt{2\pi}}\right) = \mu - \sigma\left\{\frac{\varphi(b_1) - \varphi(a_1)}{\Phi(b_1) - \Phi(a_1)}\right\} \qquad \ldots(2)$$

$$\mathsf{E}(X^2) = \lambda \int_{a_1}^{b_1} (\mu^2 + 2\mu\sigma z + \sigma^2 z^2)\varphi(z)\, dz = \lambda\mu^2 \int_{a_1}^{b_1} \varphi(z)\, dz + 2\mu\sigma\lambda \int_{a_1}^{b_1} z\varphi(z)\, dz + \lambda\sigma^2 \int_{a_1}^{b_1} z^2\varphi(z)\, dz$$

$$= \mu^2 - 2\mu\sigma\lambda[\varphi(b_1) - \varphi(a_1)] + \sigma^2\lambda \int_{a_1}^{b_1} z\frac{(ze^{-z^2/2})}{\sqrt{2\pi}}\, dz. \qquad \text{[by (2)]} \qquad \ldots(3)$$

Integrating by parts the last integral T (say) in (3), we get

$$T = \left[\frac{z(-e^{-z^2/2})}{2\pi}\right]_{a_1}^{b_1} + \int_{a_1}^{b_1} \frac{e^{-z^2/2}}{\sqrt{2\pi}}\, dz = [a_1\varphi(a_1) - b_1\varphi(b_1) + (1/\lambda)$$

Substituting the value of T in (3) yields

$$\mathsf{E}(X^2) = (\mu^2 + \sigma^2) - 2\mu\sigma\lambda[\varphi(b_1) - \varphi(a_1)] - \sigma^2\lambda[b_1\varphi(b_1) - a_1\varphi(a_1)] \qquad \ldots(4)$$

$$\therefore \; \text{Var}(X) = \mathsf{E}(X^2) - \mathsf{E}^2(X) = \sigma^2\{1 - \lambda^2[\varphi(b_1) - \varphi(a_1)]^2 - \lambda[b_1\varphi(b_1) - a_1\varphi(a_1)]\}. \qquad \ldots(5)$$

Special cases :

(i) If $b = \infty$ so $b_1 = \infty$ then $\varphi(b_1) = 0$, $b_1\,\varphi(b_1) = 0$ (L-Hospital Rule), $\Phi(b_1) = 1$.

$$\therefore \; \mathsf{E}(X) = \mu + \lambda\sigma\varphi(a_1),\; \text{Var}(X) = \sigma^2\{1 + \lambda a_1\varphi(a_1) - \lambda^2[\varphi(a_1)]^2\},\; \lambda = 1/[1 - \Phi(a_1)].$$

(ii) If $a = -\infty$ so $a_1 = -\infty$ then $\varphi(a_1) = 0$, $a_1\,\varphi(a_1) = 0$ (L-Hospital Rule), $\Phi(a_1) = 0$.

$$\therefore \; \mathsf{E}(X) = \mu - \lambda\sigma\varphi(b_1),\; \text{Var}(X) = \sigma^2\{1 - \lambda b_1\varphi(b_1) - \lambda^2[\varphi(b_1)]^2\},\; \lambda = 1/\Phi(b_1)$$

16.56. *Worked-out Problems*

***Example* 1.** Let $A \sim N(45, 4)$ and $B \sim N(44, 2.25)$ be independent. What is the prob. that

(i) two variables from A differ by 1.5 or more,

(ii) two variables from B differ by 1.5 or more,

(iii) one variable from A and one from B differ by 1.5 or more.

Solution. Let $X_i \in A$ and $Y_i \in B$, $i = 1, 2$. Then

$\mathsf{E}(X_1 - X_2) = \mathsf{E}(X_1) - \mathsf{E}(X_2) = 0$; $\text{Var}(X_1 - X_2) = \text{Var}(X_1) + \text{Var}(X_2) = 8$. $\therefore (X_1 - X_2) \sim N(0, 8)$.

$\mathsf{E}(Y_1 - Y_2) = \mathsf{E}(Y_1) - \mathsf{E}(Y_2) = 0$; $\text{Var}(Y_1 - Y_2) = \text{Var}(Y_1) + \text{Var}(Y_2) = 4.50$. $\therefore (Y_1 - Y_2) \sim N(0, 4.50)$.

Let $Z = [(X_1 - X_2) - 0]/\sqrt{8}\; T = [(Y_1 - Y_2) - 0]/\sqrt{4.5}$.

(i) $P\{(X_1 - X_2) \ge 1.5\} = P\{Z \ge 0.5303\} = 0.5000 - 0.2019 = 0.2981$.

$\therefore \; P\{|X_1 - X_2| \ge 1.5\} = 2 \times 0.2981 = 0.5962$.

(ii) $P\{(Y_1 - Y_2) \geq 1.5\} = P\{T \geq 0.71\} = 0.5000 - 0.2612 = 0.2388.$

$\therefore$ $P\{|Y_1 - Y_2| \geq 1.5\} = 2 \times 0.2388 = 0.4776$ (as above).

(iii) Let $X \in A$ and $Y \in B$; then letting $U = X - Y$, we get

$$\mathrm{E}(U) = \mathrm{E}(X) - \mathrm{E}(Y) = 45 - 44 = 1, \ \mathrm{Var}\,(U) = \mathrm{Var}(X) + \mathrm{Var}(Y) = 6.25$$

$$P\{U \leq 1.5) = P\{U^* \geq 0.5/\sqrt{6.25}\} = P\{U^* \geq 0.2\} = 0.5 - 0.0793 = 0.4207$$

Let $V = Y - X$, then $\mathrm{E}(V) = -1$, Var $(V) = 6.25$ and thus

$$P\{V \geq 1.5\} = P\{V^* \geq 2.5/\sqrt{6.25}\} = P\{V^* \geq 1\} = 0.5000 - 0.3415 = 0.1587$$

Thus $\quad P\{|X - Y| \geq 1.5\} = P(U \geq 1.5) + P(V \geq 1.5) = 0.5794.$

***Example* 2.** Let X and Y be independent $N(0, 1)$ variates. Show that $aX + bY$ and $a'X + b'Y$ are independent if $aa' + bb' = 0$. [Thus $X + Y$ and $X - Y$ are always indep.]

Solution. $M(t_1, t_2) = \mathrm{E}\{\exp[t_1(aX + bY) + t_2(a'X + b'Y)]\} = \mathrm{E}\{\exp[(at_1 + a't_2)X + (bt_1 + b't_2)Y]\}$

$= \mathrm{E}\{\exp(at_1 + a't_2)X\}.\mathrm{E}\{\exp(bt_1 + b't_2)Y\}$. [$X$ & Y are independent]

$= \exp\frac{1}{2}(at_1 + a't_2)^2 . \exp\frac{1}{2}(bt_1 + b't_2)^2 = \exp\{\frac{1}{2}(a^2 + b^2)t_1^2 + \frac{1}{2}(a'^2 + b'^2)t_2^2 + (aa' + bb')t_1 t_2\}$ ...(1)

Thus $M(t_1, t_2) = M(t_1) . M(t_2) = [\exp\frac{1}{2}(a^2 + b^2)t_1^2][\exp\frac{1}{2}(a'^2 + b'^2)t_2^2]$, iff $aa' + bb' = 0$.

Under this condition, $aX + bY \sim N(0, a^2 + b^2)$, $a'X + b'Y \sim N(a'^2 + b'^2)$.

Note. Cov $(aX + bY, a'X + b'Y) = aa' + bb'$. For bivariate Normal distribution. $\sigma_{ij} = 0 \Leftrightarrow X_i$ & X_j are independent. This proves above result without using $M(t_1, t_2)$.

***Example* 3.** Let X and Y be independent $N(0, 1)$ variates. Let $X = R\cos\Theta$, $Y = R\sin\Theta$. Show that R and Θ are independent variates.

Solution. The joint p.d.f. of indep. $N(0, 1)$ variates is

$$f(x, y) = (1/2\pi)e^{-(x^2/y^2)/2}, \quad -\infty < x, \ y < \infty.$$

Let $x = r\cos\theta$, $y = r\sin\theta$, then $\partial(x, y)/\partial(r, \theta) = r$. It follows that the joint p.d.f. of R and Θ is

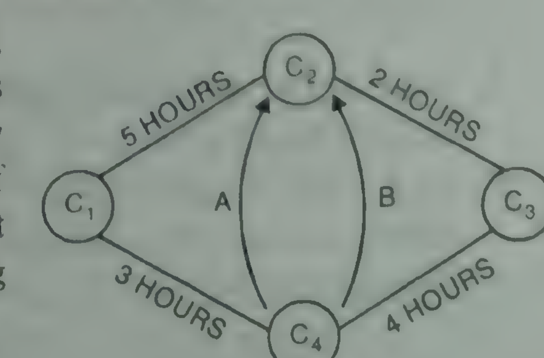
$$g(r, \theta) = f(x, y) \,|\, |\partial(x, y)/\partial(r, \theta)| = (1/2\pi)(e^{-\frac{1}{2}r^2} . r) \quad 0 \leq r < \infty . \ 0 \leq \theta \leq 2\pi$$

The factorization guarantees that R and Θ are independent distributed, with densities

$$g(r) = re^{-r^2/2}, 0 \leq r \leq \infty; \qquad h(\theta) = (1/2\pi), \ 0 \leq \theta \leq 2\pi.$$

Problems with Solutions Provided at the End of the Text

1*. The expected travel time to four cities C_i $(i = 1, 2, 3, 4)$ is indicated in the figure. The travel times for each of the routes are independently Gaussian distributed with 20% coefficient of variation (C.V.). Two cars A and B start simultaneously from C_4 to C_2, with car A going via C_1 and car B going via C_3.



(i) What is the probability that car A will arrive at C_2 within 9 hours ?

(ii) What is the probability that car A will arrive at C_2 earlier than car B ?

2*. Find the 1/1000 limits of a distribution formed by the addition of variates one from each of two independent distributions of equal size, whose parameters are

$$\mu_x = 10.5, \quad \sigma_x = 2.4, \ \mu_y = 12.5, \ \sigma_y = 3.0.$$

3*. In an examination, marks obtained by the students in Math, Physics and Chemistry are normally distributed about the means 50, 52, 48 with S.D. 15, 12, 16 respectively. Find the probability of securing a total marks of (a) 180 or above, 90 or below.

4*. If X and Y are independent values from $N(0, 9)$, what is the probability that the point (X, Y) lies between the lines $3X + Y = 5$ and $3X + Y = 10$.

5*. If $\ln_{10} X \sim N(7, 3)$ and $\ln_{10} Y \sim N(3, 1)$ are independent, find the probability of $1.202 < (X/Y) < 83180000$. [You may need $\ln_{10} 1202 = 3.80$, $\ln_{10} 8318 = 3.92$].

6*. $X_i \sim N(\mu_i, \sigma_i^2)$ $i = 1, 2, ..., n$ be independent and $Y = \sum c_i X_i$. If $(\sum c_i \mu_i)^2 = 9(\sum c_i^2 \sigma_i^2)$ find $P\{0 \le Y \le 2\sum c_i \mu_i\}$, $1 \le i \le n$.

Exercise 16(b)

1. (a) For a normal distribution, the first moment about 8 is 22 and the fourth moment about 30 is 27. Find the normal distribution.

   (b) If X, Y, Z are i.i.d. $N(0, 1)$ variates, find the mean and variance of $U = X + XY + XYZ$.

   (c) If $X \sim N(0, 1)$, find Corr (X, Y) where $Y = a = bX + cX^2$. Using Chebyshev's inequality show further that $P\{\,|Y - a - c| \le a/2\} \ge 1 - \{4(b^2 + 2c^2)'a^2\}$.

   [**Ans.** (a) $N(30, 3)$, (b) $\mu = 0, \sigma^2 = 3$, (c) $\rho = b/\sqrt{b^2 + 2c^2}$.]

2. If X is $N(\mu, \sigma^2)$, for what value of c, $E(e^X - c)^2$ is minimized. [**Ans.** $\exp[\mu + (\sigma^2/2)]$]

3. If X is $N(2, 3)$, find the p.d.f. of $Y = (1/2)X - 1$. Using tables, find $P\{Y \ge 3/2\}$.

   [**Ans.** $N(0, 3/4)$, 0.1587]

4. If $X \sim N(1, 4)$ and $Y \sim N(2, 4)$ are independent, what is the p.d.f. of $Z = X + 2Y$? [**Ans.** $N(5, 20)$]

5. If X_1, X_2, X_3, X_4 are i.i.d $N(100, 25)$ variates and $\overline{X} = \frac{1}{4}(X_1 + ... + X_4)$, find the p.d.f., mean and variance of

   (a) $4\overline{X}$, (b) $X_1 - 2X_2 + 3X_3 - 4X_4$, (c) $[\Sigma(X_i - 100)^2]/25, 1 \le i \le 4$.

   [**Ans.** $N(400, 100)$, $N(-200, 750)$, χ_4^2]

6. If X, Y, Z are indep. $N(2, 1)$, $N(3, 2)$, $N(4, 3)$ variates, find

   (a) $P(X \le Y)$, (b) $P(3X - 2Y \ge 1)$, (c) $P(X + Y \le 2Z - 4)$, (d) $P(X \le Y, Z \le 5)$.

   [**Ans.** 0.718, 0.4052, 0.3974, 0.516]

7. Let X, Y, Z be independent $N(2, 2)$, $N(3, 3)$, $N(4, 4)$ variates. Find

   (i) $P(Z + 2 \le 4X - Y \le Z + 3)$, (ii) $P(X \ge Y, Z - 3 > 0)$. [**Ans.** 0.0619, 0.2257]

8. Let $X \sim N(\mu, \sigma^2)$ and $Y \sim N(2\mu, 2\sigma^2)$ be independent. Find,

   (a) μ is $\sigma = 3$ and $P(X + 2Y \le 10) = 0.3$, (b) $\sigma = 0$ if μ and $P(4X - 3Y > 3) = 0.4$.

   (c) μ and σ if $P\{\,|2X - Y| > 10\} = 0.05$ and $P\{Y \ge 0\} = 0.9$.

   [**Ans.** $\mu = 2.945$, $\sigma = 2.057983$, (c) $\mu = 1.8826191$, $\sigma = 2.082899$]

9. Suppose $\mu \not\subset (a, b)$. At what σ will the probability $p = P(a \le X \le b)$ be the maximum ?

10. (a) Find the probability that the line joining (X, Y) to the origin makes an angle with X-axis which is less than $\pi/6$ in absolute vlaue ; X and Y indep. $N(0, 1)$ variates.
(b) Let X and Y be independent $N(0, 1)$ variates. Find the probability that (X, Y) will fall in the rectangle whose vertices are $(1.2, 0.8), (1.6, 0.6), (1.9, 1.2), (1.5, 1.4)$.
[**Ans.** (a) $p = 1/3$, (b) $p = 0.00895$]

11. If $X \sim N(6, 9)$ and $Y \sim N(7, 16)$ are indep. determine λ such that $P\{2X + Y \le \lambda\} = P\{4X - 3Y \ge k\lambda\}$, where k is a known positive real number. (Take $k = 4$).
[**Ans.** $\lambda = (3\sqrt{13} + 114\sqrt{2})/(k\sqrt{13} + 6\sqrt{2})$]

12. If $X \sim N(\mu, 9^2)$ and $Y \sim N(\mu, 12^2)$ are independent and if $P(X + 2Y \le 3) = P(2X - Y \ge 4)$, determine μ. [**Ans.** $\mu = 1.85$]

13. If $X \sim N(\mu_1, \sigma_1^2)$ and $Y \sim N(\mu_2, \sigma_2^2)$ are independent, find the relation between $\alpha, \beta, \gamma, \delta$ such that $P\{(aX + bY) \le \alpha\} = \gamma, P\{(aX + bY) \le \beta\} = \delta$. [**Ans.** $F[(\beta - \mu)/\sigma] - F[(\alpha - \mu)/\sigma] = \delta - \gamma$]

14. If $\mu_r' = E(X^r)$ is the simple moment of $N(m, \sigma^2)$, prove that

$$\mu_{r+2}' = 2m\mu_{r+1}' + (\sigma^2 - m^2)\,\mu_r' + \sigma^3\,[d\mu_r' / d\sigma].$$

15. The third decile and upper quartile of a $N(\mu, \sigma^2)$ are 56 and 63 respectively. Find μ and σ^2.
[**Ans.** $\mu = 59.0608, \sigma = 5.83$]

16. A normal population has a Coeff. of variation 3% and 75% of the population lies below 110. Find the mean and variance. [**Ans.** $\mu = 108, \sigma = 3.24$]

17. Given $N(75, 8^2)$ population, what limits will include :
(a) The middle 50% of the total frequency, (b) The middle 75% of the total frequency.
[**Ans.** (69.60, 80.40), (65.80, 84.20)]

18. If $M_X(t) = e^{2t + 32t^2}$, find $P(-2 < X < 6)$. [**Ans.** $p = 0.3820$]

19. Let X_1, X_2, X_3 be i.i.d. $N(\mu, \sigma^2)$ variates and suppose that $Y_1 = (X_1 - X_2)/\sqrt{2}$, $Y_2 = (X_1 - 2X_2 + X_3)/\sqrt{6}$, and $Y_3 = (X_1 + X_2 + X_3)/\sqrt{3}$. Show that Y_1, Y_2, Y_3 are independent normal variates. Show also $Y_1^2 + Y_2^2 = (X_1 - \overline{X})^2 + (X_2 - \overline{X})^2 + (X^3 + \overline{X})^2$, where $\overline{X} = \frac{1}{3}(X_1 + X_2 + X_3)$.

20. If $X \sim N(\mu, \sigma^2)$, find $f\{x : |x - \mu| \le k\sigma\}$. [**Ans.** $g(x) = \exp[-(x-\mu)^2/2\sigma^2]/[\sigma\sqrt{2\pi}\,.\,2\psi(k)]$]

21. Consider the p.d.f. $f_0(x) = f(x)/[1 - F(x_0)], x \ge x_0$ where

$f(x) = (2\pi\sigma^2)^{-1/2} \exp[-(x - \mu)^2/2\sigma^2],\ -\infty < \mu < \infty, \sigma \ge 0$; and $F(x_0) = P(X \le x_0)$.

Show that the first two raw moments μ_1' and μ_2' can be expressed as $\mu_1' = \mu + \lambda\sigma$, $\mu_2' = \mu^2 + \lambda\sigma\,(x_0 + \mu) + \sigma^2$ where $\lambda[1 - F(x_0)] = f[(x_0 - \mu)/\sigma]$.

22. If X is $N(0, \sigma^2)$ with density f, show that

$$E[f(X)] = \int_{-\infty}^{\infty} [f(x)]^2\,dx = \frac{1}{2\sigma\sqrt{\pi}}.$$

Hence show that if the normal distribution is grouped in intervals with total frequency N_1 and N_2 is the sum of squares of the frequencies, an estimate of σ is $N_1^2/2N_2\sqrt{\pi}$.

23. If X is $N(\mu, \sigma^2)$, show that $E(|X|) = \sigma\sqrt{2/\pi}\exp(-\mu^2/2\sigma^2) + 2\mu\,\Phi(\mu/\sigma)$ where Φ is the c.d.f. of $N(0, 1)$ variate.

24. If $X \sim N(1, 1)$, compute the covariance matrix of the random vector $Y = (X, X^2, X^2)$.

25. Determine a and b such that $f(x) = (2\pi)^{-1/2}\left[\exp(-\frac{1}{2}x^2)\right](ax^2 + bx - 3a)^2, \quad -\infty < x < \infty$ is a p.d.f. of a variate X and $E(X^4) = 3[E(X^2)]^2$. Prove that $E(X) = 0$ and $E(X^3) \neq 0$.

26. Let X and Y be independent $N(0, 1)$ variates. Find the joint p.d.f. of $Z = (X^2 + Y^2)$ and $W = X/Y$. Find the marginal densities of Z and W and show that these are independent.

27. If X and Y are independent $N(0, 1)$ variates, find the joint p.d.f. of $Z = \sqrt{(X^2 + Y^2)}$ and $W = \tan^{-1}(X/Y)$. Are Z and W independent ?

28. Let X and Y be independent $N(0, \sigma^2)$ variates. Find the probability density of

$$Z = \tan^{-1}[X/(a+Y)], \quad -\tfrac{1}{2}\pi < z < \tfrac{1}{2}\pi.$$

16-60. *Importance of Normal Distribution*

The normal distribution is of great importance in Probability theory and statistics. In nature and technology, we very often deal with distributions which are very close to normal distribution. This phenomenon is an object of investigation of the theory of Stochastic processes. Some facts of Normal Law are as under :

1. Normal law approximates the p.d.f's. of most of the commonly occurring distributions such as Binomial, Poisson, Hyp-geom, Students' t_n etc.
2. Many of the Sampling distribution such as students t, Snedecor's F, Pearsons X^2 etc. are asymptotically normal. In fact most sampling distribution tend to normality as $n \to \infty$.
3. A non-normal variate sometimes begins to exhibit normality properties under suitable transformations.
4. If Z is $N(0, 1)$, then $P\{|Z| \geq 1.96\} = 0.05$; $P\{|Z| \geq 3\} = 0.0027$. These properties of $N(0, 1)$ form the basis of "Large Sample Theory".
5. For large number (> 30) of variate observations, the sample can always be treated as normal, even though the parent population is non-normal (central limit theorem).
6. In most "Tests of Significance" the parent population is assumed to be normal.
7. 'Normal Law' is used in "Statistical Quality Control" and graduation of non-normal curves.

16-70. *Some Important Theorems*

1. *Circular Symmetry Theorem.* The variates X and Y are independent and their joint density $f_{XY}(x, y)$ has the circular symmetry (rotation invariant)

$$f_{X,Y}(x, y) = f_{X,Y}(r), \qquad r = \sqrt{x^2 + y^2}$$

Then X and Y are Normal.

Proof. We let $f_{X,Y}(r) = g(x^2 + y^2)$; use independence of X and Y to obtain ...(i)

$$f_1(x)\, f_2(y) = g(x^2 + y^2)$$

Differentiating this equation w.r.t 'x' yields

$$f'_1(x) f_2(y) = 2xg'(x^2 + y^2) \qquad \text{...(ii)}$$

Dividing Eq. (ii) by Eq. (i) gives

$$\frac{f_1'(x)}{f_1(x)} = 2x\frac{g'(x^2+y^2)}{g(x^2+y^2)} \text{ or } \frac{f_1'(x)}{2xf_1(x)} = \frac{g'(x^2+y^2)}{g(x^2+y^2)}. \qquad \text{...(iii)}$$

Differentiating (i) w.r.t 'y' gives in a similar fashion

$$\frac{f_2'(y)}{2y f_2(y)} = \frac{g'(x^2+y^2)}{g(x^2+y^2)} \qquad \text{...(iv)}$$

Eliminating g between (iii), (iv) and gives

$$[f_1'(x) / xf_1(x)] = [f_2' / yf_2(y)] \qquad \text{...(v)}$$

The L.H.S. of (v) is a function of x alone ; the R.H.S. of (v) is a function of y alone and X. Y are independent, it follows that each member of (v) must be the same constant. Hence

$$f_1'(x) / xf_1(x) = C \Rightarrow d[\ln f_1(x)] / dx = Cx \qquad \text{...(vi)}$$

Integrating : $\ln f_1(x) = \frac{1}{2}Cx^2 + b$ or $f_1(x) = ke^{Cx^2/2}$

Since $\int_{-\infty}^{\infty} f(x)\,dx = 1$; it follows that C is necessarily negative ; (If $C > 0$, the integral shall diverge) write $C = -1/\sigma^2$, get

$$f_1(x) = ke^{-x^2/2\sigma^2}.$$

This proves that X is $N(0, \sigma^2)$. From (v) and (vi) it is obvious that Y is also $N(0, \sigma^2)$. Thus X and Y are i.i.d. $N(0, \sigma^2)$ variates.

2. ***Bernstien's Theorem.*** If X_1 and X_2 are *independent* variates with the same distribution, which has finite variance and m.g.f. $M(t)$, and if $Y = X_1 + X_2$, $Z = X_1 - X_2$ are independent then all variates X_1, X_2, Y, Z are normally distributed.

Proof. We evaluate the m.g.f. of Y and Z.

$$M(t_1, t_2) = \mathsf{E}\{e^{t_1Y + t_2Z}\} = \mathsf{E}\{e^{t_1(X_1+X_2)+t_2(X_1-X_2)}\} = \mathsf{E}\{e^{(t_1+t_2)X_1} . e^{(t_1-t_2)X_2}\}$$

$$= \mathsf{E}\{e^{(t_1+t_2)X_1}\}\mathsf{E}\{e^{(t_1-t_2)X_2}\} = M_1(t_1+t_2) . M_2(t_1 - t_2) \qquad \text{...(i)}$$

Since Y and Z are independent, $M(t_1, t_2) = M(t_1, 0)\, M(0, t_2)$ and so

$$M(t_1, t_2) = \mathsf{E}(e^{t_1Y})\mathsf{E}(e^{t_2Z}) = \mathsf{E}[e^{t_1(X_1+X_2)}]\mathsf{E}[e^{t_2(X_1-X_2)}]$$

$$= \mathsf{E}(e^{t_1X_1})\mathsf{E}(e^{t_1X_2})\mathsf{E}(e^{t_2X_1})\mathsf{E}(e^{-t_2X_2}) = M(t_1)\, M(t_1)\, M(t_2)\, M(-t_2) \qquad \text{...(ii)}$$

Equating (i) and (ii); using $\theta = t_1$, $\varphi = t_2$ for brevity, we get

$$M(\theta+\varphi)\, M(\theta-\varphi) = M^2(\theta)\, M(\varphi)\, M(-\varphi) \qquad \text{...(1)}$$

We differentiate this equation w.r.t. 'φ' and obtain

$$M(\theta+\varphi)\, M(\theta-\varphi) - M(\theta+\varphi)\, M'(\theta-\varphi) = M^2(\theta)\,[M'(\varphi)\, M(-\varphi) - M(\varphi)\, M''(-\varphi)]$$

Differentiating this equation w.r.t. 'φ' we get

$$M''(\theta+\varphi)\,M(\theta-\varphi)-2M'(\theta+\varphi)\,M'(\theta-\varphi)+M(\theta-\varphi)\,M''(\theta-\varphi)$$

$$= M^2(\theta)[M''(\varphi)\,M(-\varphi)-2M'(\varphi)\,M'(-\varphi)+M(\varphi)\,M''(-\varphi)]$$

We now put $\varphi = 0$, using $M'(0) = \mu$, $M''(0) = \mu'_2$, $M(0) \equiv 1$ and get

$$M''(\theta)\,M(\theta)-2M'^2(\theta)+M(\theta)\,M''(\theta) = M^2(\theta)\,[\mu'_2-2\mu^2+\mu'_2]$$

or $$MM''-M'^2 = M^2\sigma^2,\ \text{i.e.}\ \sigma^2=(MM''-M'^2)/M^2 = d(M'/M)/d\theta$$

Integrating : $$M'/M = \sigma^2\theta + A = \sigma^2\theta+\mu. \qquad [\because M'(0)=\mu]$$

Integrating again : $$\ln M = \tfrac{1}{2}\sigma^2\theta^2+\mu\theta+B \qquad [B = 0 \text{ as } \theta = 0]$$

$$\therefore \qquad M(\theta) = e^{\mu\theta+\frac{1}{2}\sigma^2\theta^2}.$$

This is the m.g.f. of $N(\mu, \sigma^2)$. It follows that X_1 and X_2 are from the Normal distribution. Y and Z being linear combinations of X_1, X_2 are themselves normal variates.

3. *Price's Theorem.* If X is $N(\mu, \sigma^2)$, then (writing $v = \sigma^2$)

$$\frac{\partial^n}{\partial v^n}\mathrm{E}[g(X)] = \frac{1}{2^n}\int_{-\infty}^{\infty}\left\{\frac{d^{2n}}{dx^{2n}}g(x)\right\}f(x)\,dx = \frac{1}{2^n}\mathrm{E}\left\{\frac{d^{2n}}{dx^{2n}}g(X)\right\} = \frac{1}{2^n}\mathrm{E}\{D^{2n}g(X)\}$$

where $g(X)$ is an arbitrary function of X and $D = d/dx$.

Proof. By Inversion theorem, p.d.f. in terms of Ch. Function is given by

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty}e^{-itx}\,\phi(t)\,dt; \quad \phi(t)=e^{it\mu-t^2v/2} \qquad \text{...(1)}$$

$$\therefore \qquad \mathrm{E}[g(X)] = \int_{-\infty}^{\infty}g(x)\left\{\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{-itx}\,\phi(t)\,dt\right\}dx.$$

Since $\phi(t)$ depends on $v(=\sigma^2)$, differentiating above Eq. w.r.t. "v" n times we get

$$\frac{\partial^n}{\partial v^n}\mathrm{E}[g(x)] = \frac{1}{2\pi}\int_{-\infty}^{\infty}g(x)\,.\int_{-\infty}^{\infty}e^{-itx}\,\phi(t),(-\tfrac{1}{2}t^2)^n\,dt\,dx \qquad \text{...(2)}$$

because $(\partial^n\phi(t)/\partial v^n)=(-\frac{1}{2}t^2)^n\,\phi(t)$. Also, differentiating (1), $2n$ times w.r.t. x, gives

$$D^{2n}f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty}e^{-itx}(-t^2)^n\phi(t)\,dt, \qquad \{(-it)^{2n}=(-t^2)^n\} \qquad \text{...(3)}$$

$$\therefore \qquad \frac{\partial^n}{\partial v^n}\mathrm{E}\{g(X)\} = \frac{1}{2^n}\int_{-\infty}^{\infty}g(x)\,D^{2n}\,f(x)\,dx, \qquad [\text{by (2) \& (3)}] \qquad \text{...(4)}$$

Integrating by parts repeatedly and assuming that $g(x)$ does not tend to infinity too rapdidly, we obtain

$$\int_{-\infty}^{\infty}g(x)\,D^{2n}f(x)\,dx = \int_{-\infty}^{\infty}[D^{2n}\,g(x)]\,f(x)\,dx \qquad \text{...(5)}$$

The stated result follows from (4) and (5). If $\sigma^2 = v = 0$, then $X = \mu$ (= constant) ; hence

$$E[g(X)]_{v=0} = g(\mu).$$

Note. The process of integration in (5), when $n = 1$, can be easily demonstrated. Assuming $g(x) \not\to \infty$ too rapidly, e.g. if for some C and any x . $|g(x)| < C \exp(x^n)$, $n < 2$, then

$$\int_{-\infty}^{\infty} g(x)\, D^2 f\, dx = \left[g(x)\, Df\right]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} g(x)\, Dg\,.\, Df\, dx = 0 - \int_{-\infty}^{\infty} Dg\,.\, Df\, dx$$

$$= -\left[Dg\,.\, f(x)\right]_{-\infty}^{\infty} + \int_{-\infty}^{\infty} D^2 g\,.\, f(x)\, dx = \int_{-\infty}^{\infty} |D^2 g|\,.\, f(x)\, dx.$$

4. *Gaussian Law of Errors.* Under certain assumptions, the errors in the observations follow Normal probability law.

Let $X_1, X_2, \ldots, X_n$ be the n measurements of a physical quantity, these measurements are made independently and are equally reliable. If X is the *best estimate* of the quantity, the error in the individual measurements are :

$$e_1 = X_1 - X, \;\; e_2 = X_2 - X, \;\; \ldots\ldots, \; e_i = X_i - X, \;\; \ldots\ldots, \; e_n = X_n - X.$$

Obviously, $e_1 + e_2 + \ldots + e_n = (X_1 + X_2 + \ldots + X_n) - nX$.

Let us assume that the probability of making an error e is $\phi(e)$. Since the measurements X_i, $1 \le i \le n$ are made independently of each other, the probability P of simultaneous occurrence of individual errors $e_1, e_2, \ldots, e_n$ is given by

$$P = \phi(e_1)\,\phi(e_2) \ldots \phi(e_n) = \phi(X_1 - X)\,\phi(X_2 - X) \ldots \phi\,(X_n - X). \qquad \ldots(1)$$

It follows that P is the function of the *best* estimate of the physical quantity.

Assumptions and their consequences :

1. The positive and negative errors are equally likely. For precision

$$e_1 + e_2 + \ldots + e_n = 0 \;\Rightarrow\; X = (\Sigma X_i)/n = \overline{X},\; 1 \le i \le n \qquad \ldots(2)$$

i.e. the arithmetic mean X of the measurements is the best estimate of the value of the measurement. Assumption 1 requires that $\phi(-e) = \phi(e)$.

2. Small errors are more likely to occur than the large errors. Assumption 2 requires that $\phi(e)$ is a decreasing function of e, for $e > 0$.

3. Errors of great magnitude do not occur.

Assumption 3 requires that $\phi(e) \to 0$ as $e \to \pm\infty$ ...(3)

4. Best estimate X is the most probable value, i.e. the value for which P is maximum.

Derivation of the law of errors :

Assumption (4) helps to determine the form of $\phi(X)$. Taking the logarithmic differentiation in (1) and equating it to zero, we get

$$\frac{\phi'(X_1 - X)}{\phi(X_1 - X)} + \frac{\phi'(X_2 - X)}{\phi(X_2 - X)} + \ldots + \frac{\phi'(X_n - X)}{\phi(X_n - X)} = 0 \text{ i.e. } \frac{\phi'(e_1)}{\phi(e_1)} + \frac{\phi'(e_2)}{\phi(e_2)} + \ldots + \frac{\phi'(e_n)}{\phi(e_n)} = 0 \qquad \ldots(4)$$

where $e_1 + e_2 + \ldots + e_n = 0$. (Assumption No. 1). Now (2) and (4) are true for all n.

When $n = 1$, we get $e_1 = 0$ and $[(\phi'(0)/\phi(0)] = 0$, so that

$$\phi'(0) = 0. \qquad ...(5)$$

When $n = 2$, $e_2 = -e_1$ and

$$\frac{\phi'(e_1)}{\phi(e_1)} = -\frac{\phi'(e_2)}{\phi(e_2)} = -\frac{\phi'(-e_1)}{\phi(-e_1)}, \ \forall \ e_1$$

When $n = 3$, $e_1 + e_2 + e_3 = 0$ and (4) provides

$$\frac{\phi'(e_1)}{\phi(e_1)} + \frac{\phi'(e_2)}{\phi(e_2)} = \frac{\phi'[-(e_1+e_2)]}{\phi[-(e_1+e_2)]}, \ [e_3 = -(e_1+e_2)] \qquad ...(6)$$

Partial differentiation of equation (6), w.r.t. e_1 and e_2, provides

$$\frac{d}{de_1}\left[\frac{\phi'(e_1)}{\phi(e_1)}\right] = \frac{d}{de_2}\left[\frac{\phi'(e_2)}{\phi(e_2)}\right] = C, \text{ Const.}\left[\text{each} = -\frac{d}{dt}\frac{\phi'(-t)}{\phi(-t)}\right]$$

where $t = e_1 + e_2$ and $(\partial t/\partial e_i) = 1, i = 1, 2$. Since e_1 and e_2 are independent, integration of this result provides $[\phi'(e)/\phi(e)] = Ce + \lambda, (\lambda = \text{const.})$.

When $e = 0$, $\phi'(0) = 0$ by (5) and hence $\lambda = 0$. Now integrating this result in

$$\ln\phi(e) = \tfrac{1}{2}Ce^2 + \ln K \quad \text{or} \quad \phi(e) = K\exp(\tfrac{1}{2}Ce^2).$$

Since $\phi(\pm\infty) \to 0$, we must have $C < 0$; we take $C = -2h^2$, and get $\phi(e) = K\exp(-h^2e^2)$.

Now the error *certainly* lies between $-\infty$ and $+\infty$ and thus (writing y = error e)

$$1 = \int_{-\infty}^{\infty} \phi(e)\,de = K\int_{-\infty}^{\infty} e^{h^2y^2}dy = K\int_{-\infty}^{\infty} e^{-z^2/2}\frac{dz}{h\sqrt{2}} = \frac{K\sqrt{2\pi}}{h\sqrt{2}}. \qquad \left[hy = \frac{z}{\sqrt{2}}\right]$$

Thus $K = h/\sqrt{\pi}$ and so $\phi(e) = (h/\sqrt{\pi})\exp(-h^2e^2)$, $-\infty < e < \infty$.

This is Gaussian p.d.f. with precision constant h ; the standard error σ is given by $\sigma^2 = 1/2\,h^2$.

Note. To avoid e (error) with exponential 'e', write $\phi(y) = (h/\sqrt{\pi})\,e^{-h^2y^2}$, $-\infty < y < \infty$.

16-71. *Linear Transformations*

Assume $X_1, X_2, ..., X_n$ have a continuous joint distribution for which the p.d.f. is 'f' and $Y_1, Y_2, ..., Y_n$ are derived from X_i by means of a linear transformation

$$Y_i = a_{i1}X_1 + a_{i2}X_2 + ... + a_{in}X_n, \quad i = 1, 2, ..., n \qquad ...(1)$$

Let A be the $n \times n$ coefficient matrix, and write

$$A = \begin{bmatrix} a_{11} & a_{12} & ..., & a_{1n} \\ a_{21} & a_{22} & .... & a_{2n} \\ .... & .... & .... & .... \\ a_{n1} & a_{n2} & .... & a_{nn} \end{bmatrix}; \quad Y = \begin{bmatrix} Y_1 \\ Y_2 \\ ... \\ Y_n \end{bmatrix}, \ x = \begin{bmatrix} x_1 \\ x_2 \\ ... \\ x_n \end{bmatrix}, \ y = \begin{bmatrix} y_1 \\ y_2 \\ ... \\ y_n \end{bmatrix}, \ X = \begin{bmatrix} X_1 \\ X_2 \\ ... \\ X_n \end{bmatrix}$$

The linear map (1), in-matrix form, is thus given by $Y = AX$. ...(1)

Suppose det $A = |A| \neq 0$, then A^{-1} exists and Eq. (1) is a one to one transformation of the entire space R^n onto itself. At any point $(y_1, y_2, \ldots, y_n)$ of R^n, the inverse map is given by $X = A^{-1} y$. Since det $(A^{-1}) = 1/\det A$ = Jacobian of inverse map, it follows that, at any point $(y_1, y_2, \ldots, y_n)$ of R^n, the joint density $g(y_1, y_2, \ldots, y_n)$ of $Y_1, Y_2, \ldots, Y_n$ can be evaluated in the following manner :

(i) Replace component x_i in $f(x_1, x_2, \ldots, x_i, \ldots, x_n)$ by ith component of the vector $A^{-1} y$

(ii) Divide the result by | det A |.

In vector notation : $g(Y) = (|\det A|)^{-1} f(A^{-1} y)$, for $y \in R^n$. ...(2)

Linear Orthogonal Transformations :

If the square matrix A is orthogonal, then the linear map $Y = AX$, is called *linear orthogonal transformation.*

Since A is orthogonal, $A = A' = A^{-1}$, hence | det A | = 1, and transposition gives

$$Y'Y = (AX)'(AX) = X'(A'A)X = X'IX = X'X$$

$$\therefore \quad \Sigma y_i^2 = \Sigma x_i^2, \ (i = 1, 2, \ldots, n) \qquad \ldots(3)$$

16-72. Fisher's Lemma

Let $X_1, X_2, \ldots, X_n$ be independent $N(0, \sigma^2)$ variates. If X_i are transformed to new variates Y_i, $(i = 1, 2, \ldots, n)$ by linear orthogonal maps, then Y_i are also independent $N(0, \sigma^2)$ variates.

Proof. Let $Y = AX$ be the linear orthogonal map, so that $YY' = XX'$ and (det A) = 1. The joint p.d.f. of X is

$$f(x) = (\sqrt{2\pi}\sigma)^{-n} \exp\{-(\Sigma x_i^2 / 2\sigma^2), \ i = 1, 2, \ldots, n.$$

Since $\Sigma x_i^2 = \Sigma y_i^2 \ (\because XX' = YY')$ and $|J| = [\det A]^{-1} = +1$, the above equation reduces to

$$g(y) = (\sqrt{2\pi})^{-n} \exp\{-(\Sigma y_i^2 / 2\sigma^2)\}, \ i = 1, 2, \ldots, n.$$

This is the joint p.d.f. of independent $N(0, \sigma^2)$ variates $Y_1, Y_2, \ldots, Y_n$. This proves the result.

16-73. Miscellaneous Worked-out Problems

Example 1. If X is $N(\mu, \sigma^2)$, find $E(X \mid X \geq a)$ and Var $(X \mid X \geq a)$.

Solution. Let $P(X \geq a) = 1 - P(X < a) = 1 - F(a) = A$, say. We now define the *truncated* density.

$$f(x) = \frac{e^{-(x-\mu)^2/2\sigma^2}}{A\sigma\sqrt{2\pi}}, \ x \geq a; f(x) = 0, \ x < a.$$

$$\therefore \quad 1 = \int_a^\infty f(x)\,dx = \int_b^\infty \frac{e^{-z^2/2}}{A\sqrt{2\pi}}\,dz. \qquad \left[z = \frac{x-\mu}{\sigma}, \ b = \frac{a-\mu}{\sigma}\right] \qquad \ldots(1)$$

The mean M of the truncated distribution is

$$M = \int_a^\infty x f(x)\,dx = \int_b^\infty \frac{(\mu + \sigma z)e^{-z^2/2}\,dz}{A\sqrt{2\pi}} = \mu + \frac{\sigma}{A\sqrt{2\pi}}\int_b^\infty z e^{-z^2/2}\,dz = \mu + \frac{\sigma e^{-b^2/2}}{A\sqrt{2\pi}} \qquad \ldots(2)$$

$$\mathsf{E}(X^2 \mid X \ge a) = \int_a^\infty x^2 f(x)\,dx = \int_b^\infty (\mu^2 + \sigma^2 z^2 + 2\sigma\mu z)\frac{e^{-z^2/2}dz}{A\sqrt{2\pi}}, \quad [x^2 = (\mu + \sigma z)^2]$$

$$= \mu^2 + 2\mu\sigma \int_b^\infty \frac{ze^{-z^2/2}\,dz}{A\sqrt{2\pi}} + \sigma^2 \int_b^\infty \frac{z^2 e^{-z^2/2}\,dz}{A\sqrt{2\pi}}, \qquad \text{[by (1)]}$$

$$= \mu^2 + \frac{2\mu\sigma}{A\sqrt{2\pi}} e^{-b^2/2} + \frac{\sigma^2}{A\sqrt{2\pi}} \int_b^\infty z^2 e^{-z^2/2}\,dz. \qquad \ldots(3)$$

We evaluate integral involved in (3)

$$I = -\int_b^\infty z(-ze^{-z^2/2})\,dz = \left[-z.e^{-z^2/2}\right]_b^\infty + \int_a^\infty e^{-z^2/2}dz = be^{-b^2/2} + A\sqrt{2\pi}. \qquad \text{[by (1)]}$$

Substituting into (3) we get

$$\mathsf{E}(X^2 \mid X \ge a) = \mu^2 + \frac{2\mu\sigma e^{-b^2/2}}{A\sqrt{2\pi}} + \sigma^2 + \frac{\sigma^2 b e^{-b^2/2}}{A\sqrt{2\pi}}$$

$$\therefore \quad \mathrm{Var}\,(X \mid X \ge a) = \mathsf{E}(X^2 \mid X \ge a) - M^2 = \sigma^2 - \frac{\sigma^2 e^{-b^2}}{2\pi A^2} + \frac{b\sigma^2 e^{b^2/2}}{A\sqrt{2\pi}}. \qquad \left[b = \frac{a-\mu}{\sigma}\right]$$

***Example* 2.** If $X \sim N(\mu_1, \sigma_1^2)$ and $Y \sim N(\mu_2, \sigma_2^2)$ are independent, derive the distributions of $Z = (X - \mu_1)/(Y - \mu_2)$ and $W = (X - \mu_1)/\mid Y - \mu_2 \mid$.

Solution. We find Ch. Function of Z, using Double-E Rule. Note that $X_0 = (X - \mu_1) \sim N(0, \sigma_1^2)$ and $Y_0 = (Y - \mu_2) \sim N(0, \sigma_2^2)$. Now

$$\phi\,(t:Z) = \mathsf{E}\,(e^{itZ}) = \mathsf{E}\mathsf{E}\{e^{it(X_0/Y_0)} \mid Y_0\} = \mathsf{E}\,\{e^{-t^2\sigma_1^2/2Y_0^2}\} \qquad \ldots(1)$$

$$= \int_{-\infty}^\infty e^{-t^2\sigma_1^2/2y^2} \cdot \frac{e^{-y^2/2\sigma_2^2}}{\sigma_2\sqrt{2\pi}}\,dy = \frac{2}{\sigma_2\sqrt{2\pi}} \int_0^\infty e^{-y^2/2\sigma_2^2 - t^2\sigma_1^2/2y^2}\,dy \qquad \ldots(2)$$

$$= e^{-(\sigma_1/\sigma_2)|t|} \quad \therefore \left[\int_0^\infty e^{-a^2x^2 - b^2/x^2}dx = \frac{\sqrt{\pi}}{2a}e^{-2ab}, a > 0, b \ge 0; a^2 = 1/2\,\sigma_2^2, b^2 = t^2\sigma_1^2/2\right]$$

This shows that Z is Cauchy with parameter (σ_1/σ_2). Thus

$$f(z) = \frac{(b/\pi)}{b^2 + z^2}, \quad \left(b = \frac{\sigma_1}{\sigma_2}\right), \quad -\infty < z < \infty. \qquad \ldots(3)$$

Now $\phi(t : W) = \phi(t : Z)$ [since $\mid Y_0 \mid^2 = Y_0^2$ yields same results (1) and (2)]. Hence W has the same density (3) as Z has. [Another solution in Example 21-8]

***Example* 3.** $X \sim N(\mu, \sigma_1^2)$ and $Y \sim N(\mu, \sigma_2^2)$, are independent variates ($\mu < 2$) such that $P\{4X - 3Y \le 6\} + P\{5X + 12Y \ge 30\} = 1$; $P\{4X + 3Y \le 12\} + P\{5X - 12Y \le -20\} = 1$.

Determine μ and σ_1^2/σ_2^2.

Solution.

$\mathrm{Var}\,(4X \pm 3Y) = 16\sigma_1^2 + 9\sigma_2^2 = a^2$ (say) ; $\mathrm{Var}\,(5X \pm 12Y) = 25\,\sigma_1^2 + 144\sigma_2^2 = b^2$ (say) ...(1)

Using $aX + bY \sim N\,(a\mu_1 + b\mu_2, a^2\sigma_1^2 + b^2\sigma_2^2)$, we readily obtain

$(4X - 3Y) \sim N(\mu, a^2)$; $(5X + 12Y) \sim N(17\mu, b^2)$, $(4X + 3Y) \sim N(7\mu, a^2)$, $(5X - 12Y) \sim N(-7\mu, b^2)$.

The given probability statements in terms of $Z \sim N(0, 1)$ are

$$P\{Z \leq (6-\mu)/a\} + P(Z \geq (30-17\mu)/b\} = 1 \qquad \text{...(i)}$$

$$P(Z \leq (12-7\mu)/a + P\{Z \leq (7\mu-20)/b\} = 1. \qquad \text{...(ii)}$$

As $N(0, 1)$ is symmetric about $Z = 0$, $P(Z \leq A) = P(Z \geq B) \Rightarrow A = -B$. ...(iii)

(i) gives $P[Z \leq (6-\mu)/a] = P[Z \leq (30-17\mu)/b] \Rightarrow (6-\mu)/a = (30-17\mu)/b$ ...(2)

(ii) gives $P[Z \leq (12-7\mu)/a] = P[Z \geq (7\mu-20)/b] \Rightarrow (12-7\mu)/a = (20-7\mu)/b$ ...(3)

Eliminating a and b between (2) and (3) we get

$$7\mu^2 - 22\mu + 15 = 0 = (\mu-1)(7\mu-15) \Rightarrow \mu = 1,\ 15/7.$$

Since $\mu < 2$, we accept only $\mu = 1$. With $\mu = 1$, Eq. (2) or Eq. (3) provides $5b = 13a$. Now (1) yields :

$$25(25\sigma_1^2 + 144\sigma_2^2) = 169(16\sigma_1^2 + 9\sigma_2^2) \Rightarrow \sigma_2^2 = \sigma_1^2 \ \text{ or } \ (\sigma_1^2/\sigma_2^2) = 1.$$

***Example* 4.** Let X and Y be independent standard normal variates. Let $U = X + Y$, $V = X^2 + Y^2$. Show that the joint m.g.f. of U and V is

$$M(t_1, t_2) = (1-2t_2)^{-1} \exp[t_1^2/(1-2t_2)], \qquad -\infty < t_1 < \infty,\ -\infty < t_2 < \tfrac{1}{2}.$$

Hence or otherwise find Corr (U, V).

Solution.

$$M(t_1, t_2) = \mathrm{E}(e^{t_1U+t_2V}) = \mathrm{E}[e^{t_1(X+Y)+t_2(X^2+Y^2)}]$$

$$= \mathrm{E}(e^{t_1X+t_2X^2}).\mathrm{E}(e^{t_1Y+t_2Y^2}) = [\mathrm{E}(e^{t_1X+t_2X^2})]^2, \quad (X, Y \text{ are i.i.d.}) \ \text{...(1)}$$

Now

$$\mathrm{E}(e^{t_1X+t_2X^2}) = \int_{-\infty}^{\infty} \frac{e^{-x^2/2}}{\sqrt{2\pi}} e^{t_1x+t_2x^2}\, dx = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left\{\frac{(1-2t_2)x^2 - 2t_1x}{2}\right\} dx$$

Since

$$(1-2t_2)x^2 - 2t_1x = [x\sqrt{1-2t_2} - (t_1/\sqrt{1-t_2})]^2 - t_1^2/(1-t_2)$$

$$= u^2 - t_1^2/(1-t_2). \quad \left[u = x\sqrt{1-2t_2} - t_1(1-t_2)^{-12};\ dx = du/\sqrt{1-2t_2}\right]$$

$$\therefore \quad \mathrm{E}[e^{t_1X+t_2X^2}] = \frac{e^{t_1^2/2(1-2t_2)}}{\sqrt{1-2t_2}} \int_{-\infty}^{\infty} \frac{e^{-u^2/2}}{\sqrt{2\pi}}\, du = \frac{\exp[t_1^2/2(1-2t_2)]}{(1-2t_2)\,1/2}$$

Substituting in (1) provides $M(t_1, t_2) = (1-2t_2)^{-1} \exp[t_1^2/(1-2t_2)]$.

We can use cumulant generating function but direct evaluation is neater. Thus

$$\mathrm{E}(U) + \mathrm{E}(X+Y) = \mathrm{E}(X) + \mathrm{E}(Y) = 0,\ \mathrm{E}(UV) = \mathrm{E}[X^3 + XY^2 + X^2Y + Y^3] = 0$$

[X, Y are i.i.d. and odd moments are zero.]

$$\text{Cov}(U, V) = \mathrm{E}(UV) - \mathrm{E}(U)\,\mathrm{E}(V) = 0 \Rightarrow \text{Corr}(U, V) = 0.$$

Problems with Solutions Provided at the End of the Text

1*. Let $\varphi(x)$ and $\Phi(x)$ be the p.d.f. and c.d.f. of a $N(0, 1)$ variate. Let Y have a *truncated* distribution $g(y) = \varphi(y)/[\Phi(b) - \Phi(a)]$, $a < y < b$; $g(y) = 0$, elsewhere. Show that $E(Y) = [\varphi(a) - \varphi(b)] / [\Phi(a) - \Phi(b)]$.

Conversely, let $f(x)$ and $F(x)$ be the p.d.f. and c.d.f. of a continuous variates X such that $f'(x)$ exists for all x. Let

$$g(y) = f(y)/[F(b) - F(a)], \quad a \le y \le b; g(y) = 0, \text{ otherwise}$$

be the p.d.f. of a truncated variate Y. If $E(Y) = [f(a) - f(b)] / [F(b) - F(a)]$ for all real a, b prove that X is $N(0, 1)$.

2*. Find the m.g.f. of XY where X, Y are i.i.d. $N(0, 1)$ variates.

Show that the product of two independent normal variates is never a normal variate, but the product of two non-normal independent variates may be a normal variate.

3*. If the variate X_n has a c.d.f given by

$$G_n(x) = [(n-1)/n]\,\Phi(x) + n^{-1} F_n(x), \quad \forall\ x$$

where $\Phi(x)$ is the c.d.f. of a $N(0, 1)$ variate and $F_n(x)$ is a c.d.f. (of Y_n say) find the limiting distribution of X_n.

4*. If X is $N(0, \sigma^2)$, show that $Y = 1/X^2$ has p.d.f.

$$(\sigma\sqrt{2\pi})^{-1}\, y^{-3/2}\, e^{-1/2\sigma^2 y}, \quad 0 \le y \le \infty.$$

Hence or otherwise obtain the characteristic function of Y.

5*. If $X_k \sim N(0, \sigma_k^2)$, $k = 1, 2, 3$ are *independent*, then

$$Z = X_1X_2X_3 \Big/ \sqrt{X_1^2X_2^2 + X_2^2\,X_3^2 + X_3^2X_1^2}$$

is $N(0, \sigma^2)$ where $(1/\sigma) = (1/\sigma_1) + (1/\sigma_2) + (1/\sigma_3)$.

6*. (*Hard*) If $X \sim N(\mu, \sigma^2)$, find the m.g.f. and p.d.f. of $V = X^2/\sigma^2$.

7*. Find $P(A)$ the probab. of an event A if it depends on the variate $X \sim N(\mu, \sigma^2)$ and can be expressed by the formula

$$P \mid A \mid X \mid = 1 - e^{-kx}, x \ge 0,\ k > 0, \quad P \mid A \mid X \mid = 0,\ x < 0.$$

8*. Let X be a $N(0, \sigma^2)$. Let Y be another variate which is independent of X, and has $E(Y) = 0$, and is s.t. $E(Y^4) = 3\,[E(Y^2)]^2$. If $Z = X + Y$, show that there is still the relation $E(Z^4) = 3[E(Z^2)]^2$.

9*. Let $\Phi(x)$ be the c.d.f. and $f(x)$ the p.d.f. of $N(0, 1)$ variate. Prove that

(a) $\left(1 - \dfrac{1}{x^2}\right)\dfrac{f(x)}{x} \le 1 - \Phi(x) \le \dfrac{f(x)}{x}, \quad x > 0;$ (b) $\lim\limits_{x \to \infty} \dfrac{x[1 - \Phi(x)]}{f(x)} = 1.$

$\left[r(x) = \dfrac{1 - \Phi(x)}{f(x)}\right.$, is called Mill's ratio.]

Re-statement. $x^{-1}(1 - x^{-2}) \le r(x) \le x^{-1}, \quad \lim xr(x) = 1.$

10*. Let $Y_j \sim N(0, 1)$ be i.i.d. variates, $j = 1, 2, \ldots$ and define $X_n = Y_1 + Y_2 + \ldots + Y_n$. Show that

$$S_n/n^{3/2} \xrightarrow{d} N(0, \sigma^2), \text{ where } S_n = X_1 + \ldots + X_n.$$

11*. Let $Y_j \sim N(0, 1)$ be i.i.d. variates $j = 0, 1, 2, \ldots$ and define variates : $[X_n, n \geq 0]$ by

$$X_{n+1} = aX_n + Y_{n+1}, X_0 = Y_0, a < 1, \ n \geq 0.$$

Show that X_n is a Gaussian (Normal) variate and $X_n \xrightarrow{d} N(0, \sigma^2)$.

12*. Find $E(\cos X)$ and $E(\sin X)$ when X has the following *circular normal density* :

$$f(x) = [2\pi J_0(k)]^{-1} e^{k\cos(x-\beta)}, k > 0, \ 0 \leq x \leq 2\pi, \ 0 \leq \beta < \pi,$$

where $J_n(x) = \sum_{r=0}^{\infty} \frac{(x/2)^{n+2r}}{r!(n+r)!}$. $\left[J_0(x) = \sum_{r=0}^{\infty} \frac{(x/2)^{2r}}{(r!)^2}\right]$ [Bessel function of order n].

Exercise 16(c)

1. If $X \sim N(0, \sigma^2)$, show that $E(e^X) = e^{E(X^2)/2}$.
2. $X \sim N(1, 9)$ and $Y \sim N(2, 16)$ are independent. If $Z = X - Y$, write the p.d.f. of Z, state its median. S.D. and mean. Evaluate $P(Z + 1 \leq 0)$.
3. Let X and Y i.i.d. $N(0, 1)$ variates. If $Z = aX + bY + c$, what will be the distribution of Z ? Find Mean, Med. and S.D. of Z. Find $P(Z \leq 0.1)$ if $a = 1, b = -1, c = 0$.
4. Let X and Y be independently and normally distributed with a common mean. If $P(X + Y \geq 27) = P(X - Y \leq 2)$, calculate the common mean.
5. If X_i $(i = 1, 2)$ are independent and X_i is $N(i, i^2)$, find $P\{2X_1 + 4 > 1/2\ X_2 + 10\}$.
6. Let X and Y be two indep. normal r.v.s. possessing a common mean μ such that $P\{X + 2Y \leq 5\} + P\{3X + Y \leq 9\} = 1$; $P\{X - 2Y \leq 3\} + P\{Y - 3X \geq 1\} = 1$. Determine the values of μ and the ratio of variances.
7. A rocket fuel is to contain a certain percent X of a particular compound. The specification limits to X are 30 to 35 percent. The manufacturer will make a profit $p(x)$ on the fuel per gallon as follows :

   $p(x) = 10$ cents, if $30 < x < 35$. $p(x) = 5$ cents, if $35 < x < 40$ or $25 < x < 30$.
   $p(x) = -10$ cents, otherwise.
   Assuming that $X \sim N(33, 3)$, calculate expectation of $f(x)$.
8. If X is $N(0, \sigma^2)$, determine the optimal deviation E, w.r.t. the mean for which $P(a \leq X \leq b)$ will be largest $(0 < a < b)$. [E is defined by $P(|X - \mu| < E) = \frac{1}{2}$).
9. X and Y are independent normal variates with coefficient of variations η_1^{-1} and η_2^{-1} respectively. Show that $P(XY < 0) = \frac{1}{2} - 2\ \text{erf}(\eta_1)\,\text{erf}(\eta_2)$ where

$$\text{erf}(x) = \frac{1}{\sqrt{2\pi}} \int_0^x \exp\left(-\tfrac{1}{2}z^2\right) dz = \psi(x).\,.$$

10. Let X and Y be two i.i.d. $N(0, 1)$ variates. Find the joint p.d.f. of $X + Y$ and X/Y and obtain the marginal p.d.f. of $Z = X/Y$.

11. If X and Y are two i.i.d. $N(0, \sigma^2)$ variates, show that the p.d.f. of $R = \sqrt{X^2 + Y^2}$ is given by $f(r) = (r/\sigma^2) \exp(-r^2/2\sigma^2)$, $r > 0$. Find the mean and variance of R.

12. $X \sim N(0, \sigma^2)$ and Y are independent variates such that $P(Y = -q) = p$ and $P(Y = p) = q$, $p + q = 1$. Determine p and q in such a way that the moments of $Z = X + Y$ obey the relation $E(Z^4) = 3[EZ^2)]^2$.

13. If $X \sim N(\mu, v)$, show that

$$\mu_n = \frac{1}{2} n(n-1) \int_0^v \mu_{n-2}\, dv, \quad \mu_0 = 1, \ \mu_1 = 0.$$

14. If $f(x)$ and $\phi(t)$ are respectively the p.d.f and Ch. Function of a $N(0, 1)$ variate, verify

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \phi(t)\, dt \qquad \text{[Inversion Formula]}$$

15. Let $X_1, X_2, \ldots, X_n$ be i.i.d $N(\mu, \sigma^2)$ variates. Find the joint p.d.f. of $U = \sum a_i X_i$ and $V = \sum b_i Y_i$, $i = 1, 2, \ldots, n$. Hence or otherwise show that the necessary and sufficient condition that U and V are independent is that $\sum a_i b_i = 0$.

16. The frequency distribution $f(x)$ is obtained from the normal distribution $N(t) = (\sqrt{2\pi})^{-1} \exp(-t^2/2)$ by means of the equations

$$\int_1^x f(y)\, dy = \int_{-\infty}^t N(u)\, du; \ t = a \ln(x-1), \ x \geq 1.$$

If $\exp(1/a^2) = 4$, find $E(X)$ and Med X.

17. Three independent observations X_1, X_2, X_3 are given from a univariate $N(m, \sigma^2)$. Derive the joint sampling distribution of

(a) $U = X_1 - X_3$, (b) $V = X_2 - X_3$, (c) $W = X_1 + X_2 + X_3 - 3m$.

Deduce the p.d.f. of $Z = U/V$. Show that Mode $Z = 1/2$ and obtain the significant of this modal value.

18. Give a reasonable definition of $N(\mu, \sigma^2)$ if $\sigma^2 = 0$.

19. If X is $N(\mu, \sigma^2)$, find a such that $E(X \mid X \geq a) = k + \mu$, where k is a known constant.

20. A variate X has the p.d.f. $f(x) = k\varphi(x)$, $x \geq a$, $\varphi(x) = (\sqrt{2\pi})^{-1} e^{-x^2/2}$

where a is a given number, and k is a constant chosen to ensure that 'f' is a p.d.f. If $g(a) = \int_a^{\infty} \varphi(x)\, dx$, show that the mean and variance of X are respectively

$$\mu = \varphi(a)/g(a), \quad \sigma^2 = 1 + \mu(a - \mu).$$

21. Let X and Y be i.i.d. standardized variates with common Ch. function $\varphi(t)$. If $X + Y$ and $X - Y$ are independent, show that $\varphi(2t) = [\phi(t)]^3 \phi(-t)$, and deduce that X and Y are $N(0, 1)$ variates.

22. Let X and Y be independent $N(0, 1)$ variates. Show that the joint m.g.f. of $U = X + Y$ and $V = X^2 + Y^2$ is $M(t_1, t_2) = (1 - 2t_2)^{-1} \exp[t_1^2/(1 - 2t_2]$. Hence or otherwise show that U is a normal variate and that U and Y are not independent.

23. If X and Y are i.i.d. $N(0, 1)$ variates show that

(i) $(X + Y)/\sqrt{2}$ and $(X - Y)/\sqrt{2}$ are independent normal variates and find their joint distribution.

(ii) $2XY$ and $X^2 - Y^2$ have the same distribution.

24. Show that if X is conditional $N(u, v)$ and variates U and V have the joint Ch. Function $\phi(t_1, t_2)$, then the unconditional Ch. Function of X is $\phi(t_1, \frac{1}{2} t_1^2)$.

25. Let X and Y be i.i.d. $N(0, 1)$ variates. If $Z = X\sqrt{V} + Y\sqrt{1-V}$, whereas $U(0, 1)$ distribution, find the distribution of Z.

26. If X is $N(0, 1)$, find the p.d.f of $1/X^2$. If X and Y are i.i.d. $N(0, 1)$ variates, deduce that $2XY/\sqrt{X^2+Y^2}$ is $N(0, 1)$.

27. (a) If X and Y are i.i.d. $N(0, \sigma^2)$ variates, show that

$U = XY/\sqrt{X^2+Y^2}$ and $V = (X^2 - Y^2)/\sqrt{X^2+Y^2}$ are independently distributed.

(b) If $X \sim N(0, \sigma_1^2)$ and $Y \sim N(0, \sigma_2^2)$ are independent, show that $XY/\sqrt{(X^2+Y^2)}$ is $N(0, \sigma^2)$ where $\sigma^{-1} = (\sigma_1)^{-1} + (\sigma_2)^{-1}$.

28. Let $X_1, X_2, ..., X_n$ be i.i.d. $N(\mu, \sigma^2)$ variates. Let $c_1, c_2, ..., c_n$ be real constants satisfying $\Sigma c_i = 1$, $\Sigma c_i^2 = 1$. Show that $\Sigma c_i X_i$ has $N(\mu, \sigma^2)$ distribution. Show that no set of values of c's exist, which are all positive and satisfy the stated conditions.

29. A sequence of random variables (X_n) is defined by the c.d.f.'s

$$F_n(x) = \left(1 - \frac{1}{n^2-1}\right)\Phi(x) + \frac{1}{n^2-1}\Phi\left(\frac{x}{n}\right), n = 2, 3 \ldots$$ where $\Phi(x)$ is the c.d.f. of $N(0, 1)$ variate.

Show that each X_n is symmetrical about the origin, with variance 2 and that its kurtosis coefficient $Y_2 = \frac{3}{4}(n^2 - 2)$; increasing monotonically to infinity as $n \to \infty$. Also show that $F_n(x) \to \Phi(x)$ as $n \to \infty$.

30. Let $X_1 \sim N(0, 1)$ and $X_2 \sim N(1, 1/n)$ be independent variates. If a new variate Y is a result of choosing from the values $\{X_1, X_2\}$ at random with probability $1 - n^{-1}$ of choosing X_1 and n^{-1} of choosing X_2, find $P\{Y \le y\}$ and Ch. function $\phi(t : Y)$. Hence deduce that, as $n \to \infty$,

$\lim \phi(t : Y_1 + Y_2 + ... + Y_n) = \exp(-\frac{1}{2}t^2 + e^{it} - 1)$

where $Y_1, Y_2, ..., Y_n$ is the random sample from the variate Y.

LOGARITHMIC NORMAL DISTRIBUTION

For logarithmic distributions, we adopt the following way to introduce them :

$X \sim \ell n$ (Named distribution) iff $\ell n\, X \sim$ (Named distribution).

16-80. *Definition*

A positive random variable X is said to posses a ℓn-normal distribution if $\ell n\, X$ is $N(\mu, \sigma^2)$ distributed. Notationally $X \sim \ell n\text{-}N(\mu, \sigma^2)$ with parameter μ and σ^2.

Density. Here $Y \sim N(\mu, \sigma^2)$ and $Y = \ell n\, X$ or $X = e^Y$. Let $Z = (Y - \mu)/\sigma$ i.e. $N(0, 1)$, r.v. The c.d.f. of X is

$$F_X(t) = P\{X \le t\} = P\{\ell n\, X \le \ell n\, t\} = P\{Y \le \ell n\, t\} = P\{Z \le (\ell n\, t - \mu)/\sigma\}.$$

Thus, setting $z_0 = (\ell n\, t - \mu)/\sigma$.

$$F_X(t) = \int_{-\infty}^{z_0} \frac{e^{\frac{1}{2}z^2}}{\sqrt{2\pi}} dz.$$

$\therefore$ $$F'_X(t) = f_X(t) = \frac{t^{-1}}{\sigma\sqrt{2\pi}} \exp\{-\tfrac{1}{2} z_0^2\},\ t>0,\ f_x(t)=0, \quad \text{elsewhere.}$$

Moments. $\mu'_t = \mathsf{E}(X^t) = \mathsf{E}(e^{tY}) = M(t:Y) = \exp\left(\mu t + \frac{1}{2} + \sigma^2 t^2\right)$

$\therefore$ $$\mathsf{E}(X) = e^{\mu + (\sigma^2/2)},\quad \mathsf{E}(X^2) = \mu'_2 = e^{2\mu + 2\sigma^2},\quad \Rightarrow \text{Var}(X) = e^{2\mu+\sigma^2}(e^{\sigma^2} - 1).$$

Note. Some authors take $\ell n\ X \sim N(\ell n\ \zeta,\ \sigma^2)$, so that $\mu = \ell n\ \zeta$, or $\zeta = e^{\mu}$ and hence $\mu'_k = \zeta_k\ (e^{k^2\sigma^2/2}) = \zeta^k . \omega^{k^2}$, where $\omega = e^{\sigma^2/2}$. In this notation

$$\sigma^2 = \zeta^2\omega^2(\omega^2 - 1);\qquad \mu_3 = \zeta^3\omega^3(\omega^3 - 1)^2(\omega^2 + 2);\qquad \mu_4 = \zeta^4\omega^4(\omega^8 + 2\omega^6 + 3\omega^4 - 3)$$

$$\therefore\ \gamma_1 = (\omega^2 - 1)^{1/2}(\omega^2 + 2),\ \gamma_2 = (\omega^2 - 1)(\omega^6 + 3\omega^4 + 6\omega^2 + 6).$$

Since $\gamma_1 > 0$, and $\gamma_2 > 3$, the distribution is positively skew and letokurtic.

***Example* 1.** Show that the ℓn-normal distribution is positively skewed ; i.e.

mean > median > mode.

Solution. Here $f(x) = (\sqrt{2\pi}\ \sigma x)^{-1}\ e^{-\frac{1}{2}[(\ell n\, x - \mu)/\sigma]^2},\ x > 0$ ...(1)

(a) $\mathsf{E}(X^t) = \mathsf{E}(e^{tY}) = e^{\mu t + (\sigma^2 t^2/2)}$ $[Y = \ell n\, X \sim N(\mu, \sigma^2)]$. Thus $\mathsf{E}(X) = \exp\ (\mu + \frac{1}{2}\sigma^2)$.

(b) Let m be the median ; then $\int_0^m f(x)\, dx = \int_m^\infty f(x)\, dx = \frac{1}{2}$ (Put $\ell n\, x = y$)

$\therefore$ $$\int_{-\infty}^{\ell n\, m} g(y)\, dy = \int_{\ell n\, m}^{\infty} g(y)\, dy = \tfrac{1}{2} \quad [g(y) = e^{-\frac{1}{2}(y-\mu)^2/\sigma^2} / (\sqrt{2\pi}\ \sigma)]$$

This proves that the median of $N(\mu, \sigma^2)$ is $\ell n\ m$ so, $\mu = \ell n\ m$ or $m = e^{\mu}$.

(c) Take logarithm of (1) and next differentiate it to get

$$\ell n\, f(x) = -\ell n\, x - (2\sigma^2)^{-1}(\ell n\, x - \mu)^2 + \ell n\,(\sqrt{2\pi}\ \sigma)^{-1}$$

$\therefore$ $$\frac{f'(x)}{f(x)} = -\frac{1}{x} - \frac{1}{\sigma^2}\left(\frac{\ell n\, x - \mu}{x}\right) \qquad \text{...(2)}$$

Differentiating (2) again

$$\frac{f''(x)}{f(x)} - \left[\frac{f'(x)}{f(x)}\right]^2 = \frac{1}{x^2} + \frac{1}{\sigma^2 x^2}(\ell n\, x - \mu) - \frac{1}{\sigma^2 x^2} \qquad \text{...(3)}$$

$$f'(x) = 0 \Rightarrow \ell n\, x = \mu - \sigma^2 \text{ or } x = e^{\mu - \sigma^2} \qquad \text{...(4)}$$

From (3) and (4)

$$f''(x)/f(x) = -(\sigma^2/x^2)^{-1} < 0;$$ so value of x in (4) is the modal value.

Obviously : mean > median > mode $[\because e^{\mu + \frac{1}{2}\sigma^2} > e^{\mu} > e^{\mu - \sigma^2}]$

This proves that the ℓn-normal distribution is positively skewed and leptokurtic.

Example 2. If X and Y are two independent ℓn-normal variates, then XY and X/Y are also ℓn-normal variates.

Solution. Let X be ℓn-$N(\mu, \sigma^2)$ and Y be ℓn-$N(\mu', \sigma'^2)$. Let $U = XY$ and $V = X/Y$. Then $\ell n\, U = \ell n\, X + \ell n\, Y$ is the sum of two independent $N(\mu, \sigma^2)$ and $N(\mu', \sigma'^2)$ variates, it follows that U is ℓn-normal with parameters $\mu + \mu'$ and $\sigma^2 + \sigma'^2$. Further, $\ell n\, V = \ell n\, X - \ell n\, Y$; as before, it follows that V is ℓn-normal with parameters $\mu - \mu'$ and $\sigma^2 + \sigma'^2$. Hence, $u > 0$, $v > 0$.

$$f_1(u) = \frac{u^{-1}}{\sqrt{2\pi(\sigma^2+\sigma'^2)}} \exp\left[\frac{-[\ell n\, u-(\mu+\mu')]^2}{2(\sigma^2+\sigma'^2)}\right]; \; f_2(v) = \frac{v^{-1}}{\sqrt{2\pi(\sigma^2+\sigma'^2)}} \exp\left[\frac{-[\ell n\, v-(\mu-\mu')]^2}{2(\sigma^2+\sigma'^2)}\right]$$

Exercise 16(d)

1. A random variable X has the p.d.f. $f(x) = (x\sqrt{2\pi})^{-1} \exp[-\frac{1}{2}(\ell n\, x)^2, \quad x \ge 0$.
   Find the mean, mode, median, S.D. and coefficient of skewness and arrange them in order of magnitude.
2. Show that if $X \sim N(0, 1)$ is transformed by $X = \gamma + \delta\, \ell n(Y - \mu)$, the distribution of Y has
   $$\frac{\text{Mean} - \text{Mode}}{\text{Mean} - \text{Median}} = \frac{\exp[(2\delta^2)^{-1} - \exp(-\delta^{-2})}{\exp[(2\delta^2)^{-1} - 1}.$$
   Deduce that this ratio tends to 3 as $\delta \to \infty$.
3. If $\ell n\,(X - a)$ is $N(\mu, \sigma^2)$, find the distribution of X. If γ_1 is the coefficient of skewness for the distribution of X and $\sigma^2 = \ell n(1 + \theta^2)$, prove that $\theta^3 - 3\theta - \gamma_1 = 0$.
4. $X \sim N(0, 1)$ is transformed to Y by the relation
   $$X = c^{-1}[\ell n_{10}\, Y - a] \text{ with } m = e^{b^2c^2} \text{ and } b^{-1} = \ell n_{10}\, e.$$
   Show that for the transformed variable $\beta_1 = m^2(m + 3) - 4$, $\beta_2 = m^2(m^2 + 2m + 3) - 3$.
5. X is ℓn-normal variate with p.d.f.
   $$f(x\,; \mu, \sigma^2) = \frac{1}{\sigma x\sqrt{2\pi}} \exp\left\{-\frac{(\ell n\, x - \mu)^2}{2\sigma^2}\right\}, x > 0,\ \sigma > 0,\ |\mu| < \infty.$$
   Show that $e^a X^b$ has a ℓn-normal p.d.f. $f(x\,; a + b\mu, b^2\sigma^2)$. If $X_1, \ldots X_n$ are n independent observations on X, then show that $G = (X_1, X_2, \ldots, X_n)^{1/n}$ also has a ℓn-normal p.d.f. of $f(x,\ \mu. \sigma/n)$.
6. Let X, Y, Z be i.i.d ℓn-$N(4, 0.9)$ variates and let $V = XYZ$. Let $\Phi(x)$ is to be p.d.f. of $N(0, 1)$ variate. Find the numbers a, b, c such that
   $$P\{V \le e^a\} = \Phi(-1),\ P\{V \le e^b\} = 0.5,\ P\{V \le e^c\} = \Phi(1).$$
7. Let X, Y, Z be i.i.d. ℓn-$N(2, 0.9)$ variates, and let $V = XYZ$. Prove that $P\{X \ge e^2, Y \ge e^2, Z \ge e^2\} \le P\{V \le e^6\}$ and evaluate both sides. Also find $P\{V \le e^5\}$.
8. The variate X has mean m and S.D. 's' and $Y = \ell n_e X \sim N(M, S^2)$. Show that
   $$m = \exp(M + \tfrac{1}{2}S^2), 1 + (s^2/m^2) = \exp(S^2).$$
9. Given that $X_i \sim \ell n\, N(\mu_i, \sigma_i^2)$, $i = 1, 2, \ldots, n$ are independent, find $E(Y^k)$ where $Y = \prod(a_i \,.\, X_i)$.
10. Find the 5th percentile of $\ell n\, N(\mu, \sigma^2)$ when $\mu = 5.140$, $\sigma = 0.74$.

There is no grief which time does not lessen and soften. (Cicero)

Always forgive your enemies; nothing annoys them so much.

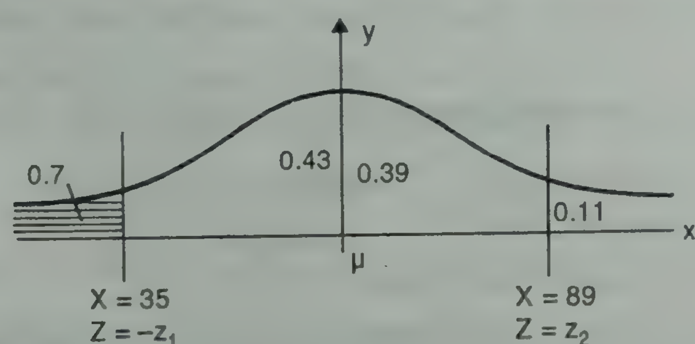
(Oscar Wilde)

Appendix : Additional Applied Problems on $N(\mu, \sigma^2)$

E

We include some worked-out problems followed by some exercises on $N(\mu, \sigma^2)$ population.

***Example* 1.** In a distribution, exactly normal, 7% of the items are under 35 and 89% are under 63. What are the mean and S.D. of the distribution ?



Solution. If $X \sim N(\mu, \sigma^2)$ is the population, setting $Z = (X - \mu)/\sigma$ we get

$$0.07 = P(X \le 35) = P\{Z \le (35-\mu)/\sigma\}; 0.89 = P(X \le 63) = P\{Z \le (63-\mu)/\sigma\}$$

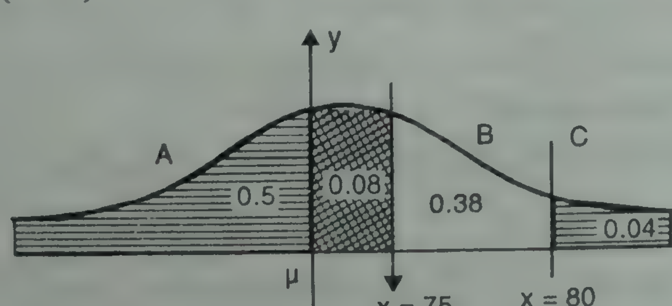
$\therefore \quad \psi(z_1) = 0.43, \ \psi(z_2) = 0.39, z_1 = (35-\mu)/\sigma, \ z_2 = (63-\mu)/\sigma$

Thus, $z_1 = 1.48, \ z_2 = 1.23$ so that $63 - \mu = (1.23)\ \sigma, \mu - 35 = (1.48)\sigma$.

Adding these Eqs. gives : $(2.71)\sigma = 28 \quad \Rightarrow \ 28/2.71 = 28 \times 0.369 = 10.332$.

$$\therefore \qquad \mu = 35 + (1.48)(10.332) = 35 + 15.29136 = 50.29.$$

***Example* 2.** If skulls are classified as A, B, C according as the length-breadth index is under 75, between 75 and 80, over 80; find approximately (assuming normality) the mean and S.D. of a series in which A are 58%, B are 38% and C are 4%, it being given that $\psi(0.20) = 0.08$, $\psi(1.75) = 0.46$.



Solution. If X denotes the length-breadth index of skulls then we are given that $P(A) = 0.58, A = \{X \ge 75\}$; $P(B) = 0.38, B = \{75 \le X \le 80\}$; $P(C) = 0.04, C = \{X \ge 80\}$.

Let $X \sim N(\mu, \sigma^2)$; and put $Z = (X - \mu)/\sigma$ i.e. $X = \mu + \sigma Z$. It is obvious that the areas give

$$P(\mu \le X \le 75) = P[0 \le Z \le (750\mu)/\sigma] = \psi[(75-\mu)/\sigma] = 0.08$$

$$P(\mu \le X \le 80) = P[0 \le Z \le (80-\mu)/\sigma] = \psi[(80-\mu)/\sigma] = 0.46$$

$$\psi[(75-\mu)/\sigma] = 0.08 \Rightarrow (75-\mu) = 75-\mu = (0.20)\,\sigma$$

$$\psi[(80-\mu)/\sigma] = 0.46 \Rightarrow (80-\mu)/\sigma = 175 \Rightarrow 80-\mu = (1.75)\,\sigma$$

Subtraction provides $1.55\,\sigma = 5 \Rightarrow \sigma = 100/31 = 3.226$

$$\therefore \quad \mu = 75-(0.20)(3.226) = 75-0.6452 = 74.355.$$

***Example* 3.** This table gives frequencies f of occurrence of a variable X between certain limits :

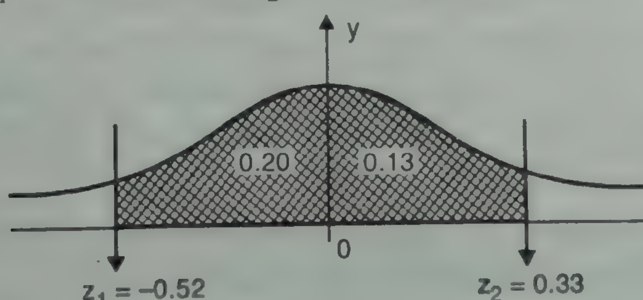
$$f(X \le 40) = 30, f(40 \le X < 50) = 33, \ f(X \ge 50) = 37.$$

The distribution is exactly normal. Find the distribution and also obtain the frequencies between $x = 50$ and $x = 60$.

Solution. Here $\sum f = 100$; so we are given, from $N(\mu, \sigma^2)$ the data

$$P(X \le 40) = 0.3, \ P(40 \le X \le 50) = 0.33, \ P(X \ge 50) = 0.37.$$

Put $Z = (X - \mu)/\sigma$, so $z_1 = (40 - \mu)/\sigma$, $z_2 = (50 - \mu)/\sigma$. The data reads



$$P(Z < z_1) = 0.3, P(z_1 < Z < z_2) = 0.33, P(Z > z_2) = 0.37.$$

From $\quad \psi(z_1) = 0.20, z_1 = -0.52 \quad$ From $\psi(z_2) = 0.13, \ z_2 = -0.33.$

$\therefore \quad 50 - \mu = \sigma(0.33)$ and $40-\mu = -\sigma(5.0) \Rightarrow 0.85\sigma = 10 \Rightarrow \sigma = 11.76$ [By Subtraction]

Also $\mu = 50-\sigma(0.33) = 50-3.88 = 46.12$. Thus, the distribution is $N(46.12, 11.76^2)$. Further

$$p = P(50 < X < 60) = P(0.33 < Z < 1.18) = \psi(1.18) - \psi(0.33) = 0.3810 - 0.1293 = 0.2517.$$

The frequencies between $X = 50$ and $X = 60$ are 25.17, i.e. 25.

***Example* 4.** A certain examination was taken by 1000 students, who are to be classified into subgroups A, B, C, D, E according to ability, range of ability to be equal in the subgroups. On the assumption that the trait measured is normally distributed, how many students should be placed in each group, given that $\psi(0.6) = 0.225$, $\psi(1.8) = 0.463$, $\psi(3.0) = 0.499$.

[$\psi(k)$ is the area under $N(0, 1)$ curve from $x = 0$ to $x = k$]

Solution. Almost entire range of $N(\mu, \sigma^2)$ is from $\mu - 3\sigma$ to $\mu + 3\sigma$, it follows that the total range is $(\mu + 3\sigma) - (\mu - 3\sigma) = 6\sigma$. This range is to be divided into five equal subgroups are

$(\mu - 3\sigma, \mu - 1.8\sigma)$, $(\mu - 1.8\sigma, \mu - 0.6\sigma)$, $(\mu - 0.6\sigma, \mu + 0.6\sigma)$, $(\mu + 0.6\sigma, \mu + 1.8\sigma)$, $(\mu + 1.8\sigma, \mu + 3\sigma)$

Or measured in standard units ($\mu = 0$, $\sigma = 1$), these are

$$(-3, -1.8), (-1.8, -0.6), (-0.6, 0.6), (0.6, 1.8), (1.8, 3)$$

If $f(a, b)$ denotes the frequency of group in the range (a, b), then,

$$f[1.8, 3] = 1000 \times [\psi(3) - \psi(1.8)] = 1000 \times (0.499 - 0.463) = 36$$

$$f[0.6, 1.8] = 1000[\psi(1.8) - \psi(0.6)] = 463 - 225 = 238;$$

$$f[-0.6, 0.6] = 1000[\psi(0.6) + \psi(0.61)] = 450.$$

Thus, the number of students in the groups are 36, 238, 450, 238, 36 and these figures total to 998. The rest of the two students can be put into the extreme groups $(\mu - 3\sigma, \mu - 1.8\sigma)$ and $(\mu + 1.8\sigma, \mu + 3\sigma)$ because these intervals are open. We can thus keep 37, 238, 450, 238, 37 students into the group A, B, C, D, E respectively.

Problems with Solutions Provided at the End of the Text

1*. Suppose diameters of shafts are normally distributed with mean 10 cm and S.D. 0.1 cm. If the shaft must meet the specification that its diameter fall between 9.9 and 10.2 cm, what proportion of shafts will meet specification ?

2*. A manufacturer making beds finds that the average height of men is 1 m. 72 cm with a S.D. of 8.5 cm. What length (to the nearest cm) must this firm make its beds, in order that no more than 5% of men, find themselves too long for the beds manufactured ?

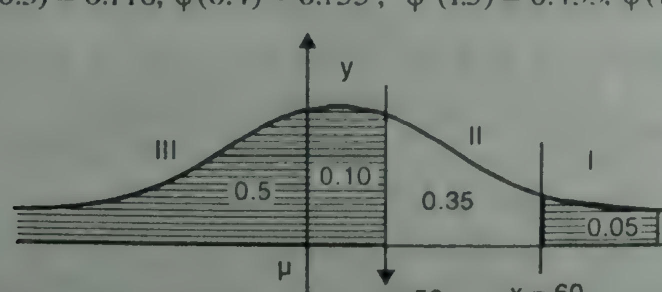
3*. Assuming the mean height of soldiers to be 68.22 inches, variance 10.8 (inches)^2, find how many soldiers in a regiment of 1000 would you expect to be over 6 feet tall. [Given : area under $N(0, 1)$ curve between $z = 0$ and $z = 0.35$ is 0.1368 and between $z = 0$ and $z = 1.15$, it is 0.3746].

4*. In $N(15, 3.5^2)$ population, it is known that 647 observations exceed 16.25. What is the total number of observations in the population ?

5*. The S.D. of a certain group of 1000 high school grades was 11% and their mean grade 78%. Assuming the distribution to be normal find (i) how many grades were above 90%, (ii) what was the highest grade of the lowest 10, (iii) what was the semi-interquartile range and (iv) within what limit did the middle 900 lie ?

6*. One thousand candidates in an examination were grouped into three classes I, II and III in descending order of merit. The number in the first two classes were 50 and 350 respectively. The highest and lowest marks in class II were 60 and 50 respectively. Assuming normal distribution, find the average, S.D. and the number of candidates obtaining marks between 43 and 53. Areas under $N(0, 1)$ curve are

$\psi(0.2) = 0.079$, $\psi(0.3) = 0.118$, $\psi(0.4) = 0.155$; $\psi(1.5) = 0.433$, $\psi(1.6) = 0.445$, $\psi(1.7) = 0.455$



7*. The marks obtained by a number of students for a certain subject are assumed to be approximately normally distributed with mean 65.2 and a S.D of 5. If 5 students are taken at random from this set, what is the probability that 3 of them will have marks over 75 ?

8*. Women's shoes are manufactured in sizes 2, 3, 4, ..., 8 ; the length advancing in steps by 1/4 in. Size 5 is suitable for a foot of length ranging from $9\frac{1}{4}$ inch to $9\frac{1}{2}$ inch. If the length of women's feet are normally distributed with mean 9.4 inch and S.D. 0.25 in., how many pairs of shoes of each size will be required, out of every 10,000 pairs manufactured ? [Assume that all women with feet less than $8\frac{3}{4}$ in length will be content with size 2 and women requiring a size larger than 8 are not catered for].

Exercises

1. (a) In a normal distribution, 31% of the items are under 45 and 8% are over 64. Find the mean and S.D. of the distribution. **[Ans.** 49.96, 9.995]

   (b) Of a large group of men, 5% are under 60″ in height and 40% are between 60″ and 65″. Assuming a normal distribution, show that the mean height is 65.5 and S.D. is 3.3″.
2. The width of a slot on a forging is $N(0{,}9'', [0.004'']^2)$. The specifications are 0.900 ± 0.005 inch. What percentage of forgings will be defective ?
3. Given that X is normally distributed with mean 10, and $P(X > 12) = 0.1587$. Show that the probability that $X \in (9, 11)$ is 0.3830. You may use $\Phi(1) = 0.8413$ and $\Phi(-\frac{1}{2}) = 0.3085$.
4. A dealer sells an item for Re 1. If the weight $W \sim N(\mu, 1)$ of the item is less than w_0, it is dead loss. The cost c per item is $c = a + bW$, $(a > 0, b > 0)$. Determine μ so that expected profit is maximum $(\mu > w_0)$.
5. A minimum height is to be prescribed for eligibility to Government service such that 60% of the young men will have a fair chance of coming upto that standard. The heights of youngmen are normally distributed with mean 60.6″ and S.D. 2.55″. Show that the minimum specification is 59.96″.
6. The time when a country bus passes a certain point is distributed normally with a mean 9.25 A.M. and a S.D. of 3 min. What is the least time one could arrive at this point and still have a probability of 0.99 of catching the bus ?
7. Torch batteries of a certain make have an average life of 50 hrs with a S.D. of 3 hrs. How many batteries in an order of 1000 may be expected to last (a) longer than 55 hrs, (b) less than 44 hrs ? Find approximately within what time 30% of the batteries may be expected to fail. **[Ans.** 48, 23 ; time : 47.45 hours]
8. Five thousand candidates appeared in a certain examination paper carrying a maximum of 100 marks. It was found that the marks were distributed as $N[39.5, (12.5)^2]$. Determine approximately the number of candidates who secured a first class for which a minimum of 60 marks is necessary. You may use the following table :

   $Z = 1.5, 1.6, 1.7, 1.8$ Corresponding $A = 0.9332, 0.9452, 0.9554, 0.9641$

   Here A is the proportion of the *whole* area of the normal curve lying to the **left** of the ordinate at z. **[Ans.** 272]
9. In an examination it is laid down that a student passes if he secures 30% or more marks. He is placed in the first, second or third division according as he secures 60% or more, between

45% and 60% and between 30% and 45% marks respectively. He gets a distinction in case he secures 80% or more marks. It is observed from the results that 10% of the students failed in the examination, whereas 5% of them obtained distinction. Show that the percentage of the students placed in the second division is 34%. Assume normal distribution of marks.

10. In a University examination of a particular year, 60% of the students failed when mean of marks was 50% and S.D. 5%. What were the minimum pass marks ? University decided to relax the conditions of passing by lowering the pass marks to show its result 70%. Find the minimum marks for a student to pass, supposing the marks to the normally distributed and no change in the performance of students takes place. [**Ans.** 47.38]

11. The average percentage of marks of candidates in an examination is 42 with a S.D. of 10. The minimum for a pass is 50% and if 1000 candidates appear for the examination, how many can be expected to pass assuming that marks are distributed normally ? If it is required that double that number should pass, what should be the average percentage of pass marks ?

12. A certain number of articles manufactured in one batch were classified into three categories according to a particular characteristic being less than 50, between 50 and 60, or greater than 60. If this particular characteristic is known to be normally distributed, determine the mean and S.D. for this batch if 60%, 35% and 50% were found in these categories. Given that if erf (0.25) = 0.1 and erf (1.66) = 0.15).

13. Marks obtained in Statistics in a certain examination are found to be normally distributed. If 12.5% of the candidates obtain 60% or more marks, 30% obtain less than 30 marks, find the mean number of marks obtained by the candidates $\Phi(z)$ is the c.d.f. of $N(0, 1)$.

z	:	0.27	0.28	0.29	1.14	1.15	1.16
$\Phi(z)$	:	0.6064	0.6102	0.6104	0.8727	0.8749	0.8770

14. The mean I.Q. (intelligence quotient) of a group of 225 children is 102 and S.D. is 15 with reliability (of the test) 0.91.

(i) What proportion of this group is expected to have I.Q's greater than 132 ?

(ii) How often will the mean I.Q.'s of the group fall as low as 100 ? What is the probability that the mean I.Q. of the group is above 100 ?

15. A nurse lives at A and works at C. She starts work at 8 a.m. She always uses train from A to B which generally reaches B at 7 : 40 a.m. Buses from B leave for C every 15 minutes and the bus that leaves at 7 : 45 a.m. generally arrives C at 7 : 56 a.m. The train on average gets delays of 2 minutes and has a S.D. 4 minutes. The bus always leaves on time, but reaches C, on average, after a 2-minute delay and a S.D. of 3 minutes. Show that $P\{\text{Nurse arrives late}\} = 0.40$. The doctor (nurse's employer) drives to C, he leaves at 7 : 45 a.m. and driving time to C is $N(12, 2^2)$ in minutes. Show that

$P\{\text{Nurse and doctor both arrive late}\} = 0.028$.

$P\{\text{Doctor arrives earlier than the nurse}\} = 0.61$.

16. In a certain examination, the percentage of passes and distinctions were 45 and 9 respectively. Estimate the average marks obtained by the candidates, the minimum pass and distinction marks being 40 and 74 respectively. Assume the distribution of marks to be normal. Also determine what would have been the minimum of qualifying marks for admission to a re-examination of the failed candidates, had it been desired that the best 25% of them should be given another opportunity of being examined.

17. Plugs and Sockets are manufactured by independent machines. Let P and S denote the diameters in millimeters of a plug and a socket respectively. Socket fits well to the plug if S is in the

range $(P + 0.2, P + 0.4)$. Assume that P is $N(1, 0.001)$ and S is $N(1.3, 0.0015)$. What percentage of plugs and sockets fit well ?

18. (a) Marks secured by students in sections A and B of a class are independently normally distributed with means 50 and 60 respectively and variances 10 and 6 respectively. What is the probability that a randomly chosen student from section A scores more than a randomly chosen student from section B ? What percentage of students are expected to secure first division. (i.e. 60 marks or more) in section A ? Write down your results in terms of the standard normal c.d.f. $\Phi(z)$.

(b) Articles are being manufactured with a dimension whose mean value is 2.522 in. with S.D. 0.005 in. Tolerance limits are 2.510 in. and 2.530 in. Find what percentage of articles will fail to meet the tolerances if the distribution is normal.

19. Steel rods are manufactured to be 3 inches in diameter but they are acceptable if they are inside the limits 2.99 inches and 3.01 inches. It is observed that about 5% are rejected over-size and 5% are rejected under-size. Assuming the diameters are normally distributed, show that S.D. of the distribution is 1/165. Hence calculate what proportions of rejects would be if the permissible limits were widened to 2.985 and 3.015 inches. [**Ans.** 1.34%]

20. A company uses many thousands of electric lamps annually burning day and night. Assume that under such conditions the life of a lamp may be regarded as $N[50 \text{ days}, (19)^2 \text{ days}]$. On January 1, 1972, the company put 5000 new lamps into service. How many expected to be replaced by (i) February 1, (ii) April 1 ? The lamps are put into operation at about the same time of the day.

21. In an examination, the mean and S.D. of marks in mathematics and chemistry are shown in the adjoining table. Assuming the marks in two subjects to be independent normal variates, show that the probability that a student scores total marks lying between 100 and 130 is 0.3635. [Full marks in each subject are 100]. Given that erf (0.28) = 0.103, erf(1.94) = 0.4738. [erf $(z) = \psi(z)$.]

	Mean	S.D.
Maths	45	10
Chem.	50	15

22. Two pieces AB and CD are assembled together in a mass production, assembly. Lengths AB and CD are normally distributed. The average length of CD is 2.5 inches with S.D. of 0.13 inches. Both dimensions are normally distributed and the assembly is at random. Find the approximate probability the assembly will have a combined length of 6.05 inches or more.

23. Screws are rejected if the length is less than ℓ_1 and greater than ℓ_2. From a batch of 128 screws produced by a certain machine 28 were found to be less than ℓ_1 in length and 8 were longer than ℓ_2. Assuming the lengths to be approximately normally distributed, estimate the mean length and S.D. of the screws. Can you estimate the average length of screws which are accepted ?

24. In a certain human population, head index i is assumed to be normally distributed. There are 58% dolichocephatics ($i \le 75$), 38% mesocephatics ($75 \le i \le 80$) and 4% brachycephatics ($i > 80$). Find the mean and S.D. of i. If F is the standard cumulative normal distribution, then it is known that $F(0.202) = 0.58$, $F(1.75) = 0.95$.

25. A manufacturing company produces rods. It is desired to have the length of a rod 4 cms, but infact the length has normal distribution with mean = 4.01 cms and S.D. = 0.03 cms. Each rod costs 12 paise to make and may be used immediately if its length lies between 3.98 cms and 4.02 cms. If the length of the rod is less than 3.98 cms, the rod is useless but has the scrap

value of 3 paise. If its length exceeds 4.02 cms, it may be shortened and used at a further cost of 4 paise. Find the cost per usable rod.

26. The mean yield for one-acre plot is 662 kilos with a S.D. 32 kilos. Assuming normal distribution, how many one-acre plots in a batch of 1000 plots would you expect to have yield (i) over 700 kilos, (ii) below 650 kilos, (iii) what is the lowest yield of the best 100 plots ?

27. The local authorities in a certain city install 10,000 electric lamps in the streets of the city. If these lamps have an average life of 1000 burning hours with a S.D. of 200 hours, assuming normality, how many lamps might be expected to fail (i) in the first 800 burning hours ? (ii) between 800 and 1200 burning hours ? After what period of burning hours would you expect that (a) 10% of the lamps would fail ? (b) 10% of the lamps would be still burning ? [In $N(0, 1)$, $\psi(1) = 0.34134$ and 80% area falls between the limits $z = \pm 1.28$].

28. The local authorities in a certain city installed 2000 lamps in the streets of the city. The lamps have an average life of 1000 burning hours with a S.D. of 200 hours.

(a) How many lamps might be expected to fail in the first 700 burning hours ?

(b) After what periods of burning hours would we expect that

(i) 10% of the lamps would have failed ? (ii) 10% of the lamps would be still burning ?

Assume that the lives of the lamps are normally distributed. Given $\Phi(1.50) = 0.933$, $\Phi(1.28) = 0.900$.

29. An establishment uses 1000 bulbs which are kept burning approximately 4 hours everyday. Past experience of cost incurred in replacing burnt-out bulbs has shown that it is profitable to replace all the 1000 bulbs, whether some of them are burnt out or not, in 4 months (about 120 days). Bulbs burning out during this period are not replaced till the end, inconvenience regretted. Assuming the life in days of a bulb is normally distributed $N(450, 30^2)$, find the expected number of hours for which a bulb is dead at any point during the four month period.

A few observations and much reasoning lead to error, many observations and a little reasoning lead to truth. [Alexis Carrol]

We know a thing when we understand it. (George Berkeley)

Central Limit Theorem. Normal Approximations

17

Notation. Let $X_1, X_2, \ldots, X_n$ be i.i.d random variables from a distribution (population) which has a finite mean $\mu(= \mathsf{E}(X)$ and variance σ^2 and any proabability law. We record this statement simply as

$$(X_1, X_2, \ldots, X_n) \underbrace{\quad}_{\text{i.i.d.}} (\mu, \sigma^2)$$

Of course, when the p.d.f. (or p.m.f.) of the distribution is known, we modify the notation, as usual and specify the parameters as well. For example,

$$X_i \sim \text{bin}\,(n, p),\ X_j \sim N(\mu, \sigma^2),\ X_k \sim \text{gam}\,(\alpha, \lambda), \ldots$$

17-10. *Definition of Central Limit Theorem (C.L.T.)*

A sequence of jointly distributed random variables $X_1, X_2, \ldots, X_n$ with finite means and variances, $0 < \text{Var}\,(X_i) < \infty$, is said to obey the central limit theorem (CLT) iff the standardized Sum S_n^*.

$$S_n^* = \frac{S_n - \mathsf{E}(S_n)}{\sigma(S_n)} \text{ converges in distribution to } N(0, 1)\text{-variate.} \qquad \left[S_n = \sum_{i=1}^{n} X_i = n\overline{X}_n\right]$$

Remark. By Levy's Continuity Theorem, $\{X_n\}$ obeys C.L.T. iff, $\forall\ t$

$$M(it : S_n^*) \to \exp\,(-t^2/2), \text{ as } n \to \infty.$$

Note. Consider independent X_i's, each with mean μ and variance σ^2. Then

$$S_n^* = \frac{S_n - n\mu}{\sigma\sqrt{n}} = \frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}} \qquad \text{[Standardized sample sum or mean]}$$

17-11. *Lindeberg-Levy C.L.T.*

Let $X_1, X_2, \ldots, X_n$ be i.i.d. variates, with mean μ and variance σ^2, $0 < \sigma^2 < \infty$. Let

$$F_n(x) = P(S_n^* \le x), \qquad \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-\frac{t^2}{2}\right) dt.$$

Then $F_n(x) \to \Phi(x)$ as $n \to \infty$ for all $x \in R$.

Equivalently : $\left|[P(S_n^* \le z) - \Phi(z)\right| = \left|[P(X_n^* \le z) - \Phi(z)\right| < \varepsilon$.

Proof. Firstly we observe that

$$S_n^* = \frac{S_n - \mathsf{E}(S_n)}{\sigma(S_n)} = \frac{S_n - n\mu}{\sigma\sqrt{n}} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left(\frac{X_i - \mu}{\sigma}\right) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} Y_i$$

where $Y_i = [(X_i - \mu)\sigma]$, are i.i.d standardized variates with $\mathsf{E}(Y_i) = 0$, Var $(Y_i) = 1$. The M.G.F. of S_n^* is given by

$$M(t:S_n^*) = M[t:(\Sigma Y_i)/\sqrt{n}] = M[t/\sqrt{n}:\Sigma Y_i] = [M(t/\sqrt{n}:Y_i)]^n. \quad ...(1)$$

Now Taylor's series for $M_Y(t/\sqrt{n}) \equiv M(t')$ is

$$M(t') = \sum_{r=0}^{\infty} \frac{(t')^r}{r!} M^{(r)}(0) = 1 + \frac{(t^2/n)}{2!} + \frac{(t/\sqrt{n})^3}{3!}\frac{\mu_3}{\sigma^3} + \ldots \qquad \left[M_{(0)}^{(r)} = \frac{\mu_r}{\sigma^r}, \mu_1 = 0\right]$$

$$\therefore \qquad [M(t')]^n = \left[1 + \frac{t^2}{2n} + \mathrm{O}\left(\frac{t^2}{n}\right)\right]^n$$

So $\displaystyle\lim_{n\to\infty} M(t:S_n^*) = \lim_{n\to\infty}\left[1 + \frac{t^2}{2n} + O\left(\frac{t^2}{n}\right)\right]^n = \exp\left(\frac{t^2}{2}\right),$ (by Euler's limit Theorem)

As $e^{t^2/2}$ is the m.g.f of $N(0, 1)$ distribution, it follows by Levy's Continuity Theorem that $S^*_n \to N(0, 1)$ in distribution. Equivalently, $\overline{X}_n \to N(0, 1)$ in distribution.

In other words $\langle X_n \rangle$ follows C.L.T. Further, since $\Phi(x)$ is continuous c.d.f. so by Polya's theorem, $F_n(x) \to \Phi(x)$ uniformly $\forall\, x \in \mathbb{R}$.

Cor. Let $X_1, X_2, \ldots, X_n$ be i.i.d. *Standardized* variates such that $n^{-1/2} S_n$ has the same distribution, $\forall\, n = 1, 2, \ldots$ Then, the distribution of X_i must be $N(0, 1)$.

Proof.
$$S_n^* = \frac{S_n - \mathrm{E}(S_n)}{\sigma(S_n)} = \frac{S_n - 0}{\sqrt{n}} = \frac{S_n}{\sqrt{n}}. \qquad [F_n(x) = \text{c.d.f of } S_n^*]$$

$$\therefore \qquad \lim_{n\to\infty} F_n(x) = \lim_{n\to\infty} P\{S_n/\sqrt{n} \le x\} = \Phi(x) \qquad \text{[by C.L.T]}$$

Since $(S_n/\sqrt{n})$ has the same distribution for all n, we must have $P\{n^{-1/2} S_n \le x\} = F(x)$ for each n. It follows that we must have $F(x) = \Phi(x)$, *i.e.* $X_i \sim N(0, 1)$ for each n.

17-12. *De-Moivre's Laplace C.L.T.*

Let $X_1, X_2, \ldots, X_n$ be i.i.d. Bernoulli random variables with common probability p of success. If $X = X_1 + X_2 + \ldots + X_n$ $(= S_n)$, then

$$\lim_{n\to\infty} P\left\{\frac{X - np}{\sqrt{npq}} \le x\right\} = \Phi(x).$$

Proof. Here $\mathrm{E}(X_i) = p, \mathrm{E}(X_i^2) = p, \mathrm{Var}(X_i) = \mathrm{E}(X_i^2) - \mathrm{E}^2(X_i) = pq,$ so $\mathrm{E}(S_n) = np, \mathrm{Var}\, S_n = npq.$

Hence by Lindeberg-Levy C.L.T.

$$\Phi(x) = \lim_{n\to\infty} P\{S_n^* \le n\} = \lim_{n\to\infty} P\left\{\frac{X - np}{\sqrt{npq}} < x\right\}. \qquad \left[S_n^* = \frac{S_n - \mathrm{E}(S_n)}{\sigma(S_n)}\right]$$

This completes the proof. Consult also §17-51.

17-13. *Stirling's Approximation*

For large values of n, $\qquad n\,! \approx \sqrt{2\pi}\, e^{-n}\, n^{n+(1/2)}.$

Proof. Let $X_1, X_2, \ldots X_n$ be i.i.d. Pois (1) variates. Then $S_n = X_1 + X_2 + \ldots + X_n$ is Pois (n), by Addition Theorem. Here $\mathrm{E}(S_n) = n = \mathrm{Var}\,(S_n)$. Consequently,

$$P\{S_n = n\} = e^{-n} n^n / n! \qquad ...(1)$$

We evaluate $P\{S_n = n\}$ by C.L.T. as well. As S_n is discrete, we apply continuity correction as well. Now.

$$P\{S_n = n\} = P\{n - \tfrac{1}{2} \le S_n \le n + \tfrac{1}{2}\}$$

$$= P\left\{\frac{(n-1/2)-n}{\sqrt{n}} \le \frac{S_n - n}{\sqrt{n}} \le \frac{(n+1/2)-n}{\sqrt{n}}\right\},$$

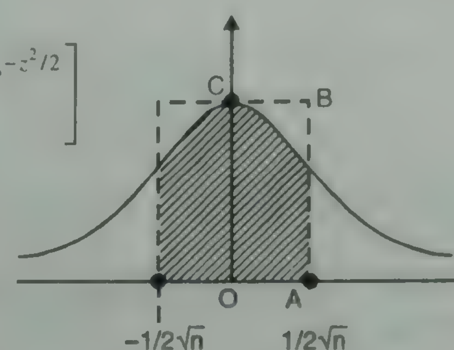
$$= P\left\{-\frac{1}{2\sqrt{n}} \le S_n^* \le \frac{1}{2\sqrt{n}}\right\}, \qquad \left[S_n^* = \frac{(S_n - n)}{\sqrt{n}}\right]$$

By C.L.T., $S_n^* = Z \sim N(0, 1)$, hence

$$P\{S_n = n\} = 2\int_0^{1/2\sqrt{n}} f(z)\,dz = 2 \times \text{Area of Rect. } OABC^\dagger$$

$$= 2 \times \left(\frac{1}{2\sqrt{n}} \times \frac{1}{\sqrt{2\pi}}\right) \quad \ldots(2) \quad \left[\begin{array}{l} f(z) = (1/\sqrt{2\pi})\ e^{-z^2/2} \\ f(0) = 1/\sqrt{2\pi} \end{array}\right]$$

$$= 1/\sqrt{2n\pi}$$



Equating (1) and (2) gives

$$e^{-n}n^n / n! \approx 1/\sqrt{2n\pi}$$

This gives $n! \approx \sqrt{2\pi}\, e^{-n}\,.n^{n+(1/2)}$

***Example* 1.** Let $X_1, X_2, \ldots, X_n$ be a sequence of independent Bernoulli variates :

$$P(X_i = 1) = p,\ P(X_i = 0) = q,\ p + q = 1, \text{ Let } X = (\Sigma X_i)/n,\ \ 1 \le i \le n.$$

Verify C.L.T. by showing that X tends to be normal as $n \to \infty$.

Solution. $(\mathrm{E}X_i) = 1.p + 0.q = p$. $\mathrm{E}(X_i^2) = 1^2.p + 0^2.q = p$, so $\mathrm{Var}(X_i) = \mathrm{E}(X_i^2) - \mathrm{E}^2(X_i) = p - p^2 = pq$.

$$\mathrm{E}(X) = [\Sigma \mathrm{E}(X_i)]/n = np/n,\ \mathrm{Var}(X) = (1/n^2)[\Sigma\ \mathrm{Var}(X_i)] = npq/n^2 = pq/n = \sigma^2 \text{ (say).}$$

$$M(t : X_i) = \mathrm{E}(e^{tX_i}) = pe^t + q;\ M(t : \bar{X}) = M(t/n : \Sigma X_i) = [M(t/n : X_i)]^n = [q + pe^{t/n}]^n. \quad \ldots(1)$$

As per C.L.T. $X^* = (X - p)/\sqrt{pq/n}$ is approximately Normal if n is large. To verify it,

$$M(t : X^*) = M[t; (X-p)/\sigma] = e^{-pt/\sigma} M[t/\sigma) : X] = e^{-pt/\sigma}[q + pe^{t/\sigma n}]^n.$$

$$= (pe^{qt/n\sigma} + qe^{-pt/n\sigma})^n = \left[1 + \frac{t^2}{2n} + O\left(\frac{1}{n^{3/2}}\right)\right]^n$$

$$\therefore \quad \lim_{n\to\infty} M(t : X^*) = \lim_{n\to\infty}\left[1 + \frac{t^2}{2n} + O\left(\frac{1}{n^{3/2}}\right)\right]^n = e^{\frac{1}{2}t^2}$$

This shows that the asymptotic distribution of X^* is $N(0, 1)$

†Area of infinitesimal rectangle amounts to Taylor's approximation to the function $f(z)$.

Note. $M(t:X) = [1+p(e^{t/n}-1)]^n = [1+(pt/n)+O(1/n)]^n$. As $n \to \infty$, $M(t:X) = e^{pt}$ (m.g.f. of a degenerate r.v.). However, X is not degenerate since Var $X \neq 0$.

***Example* 2.** Prove the C.L.T. for a sequence of independent variates $\langle X_n \rangle$ having the distribution (*i*) $U(-a, a)$. (ii) Expo (λ).

Solution. (i) Here $E(X_i) = (a-a)/2 = 0$, Var $X_i = (b-a)^2/12 = a^2/3$. Let $S_n = X_1 + \ldots + X_n$

Then $E(S_n) = 0$, $B_n =$ Var $(S_n) = n$ Var $(X_1) = na^2/3$ so $S_n^* = (S_n - 0)\sqrt{3}/a\sqrt{n}$. Now

$M(t:S_n^*) = M\left(t:\sqrt{3}\,S_n/a\sqrt{n}\right) = M(\sqrt{3}t/a\sqrt{n}:S_n) = [M(\sqrt{3}\,t/a\sqrt{n}:X_i)]^n$

Since $$M(t:X_i) = E(e^{tX_i}) = \int_{-a}^{a}\frac{e^{tx}}{2a}dx = \frac{e^{at}-e^{-at}}{2at} = \frac{\sinh at}{at}.$$

$$\therefore \quad M(t:S_n^*) = \left[\frac{\sinh(t\sqrt{3}/n)}{t\sqrt{3}/n}\right] = \left[1+\frac{3t^2}{3!\,n}+0\left(\frac{1}{n}\right)\right]^n$$

So $$\lim_{n\to\infty} M(t:S_n^*) = \lim_{n\to\infty}\left[1+\frac{t^2}{2n}+0\left(\frac{1}{n}\right)\right]^n = e^{t^2/2} = [\text{m.g.f. } N(0,1)] \text{ by Euler's limit.}$$

(ii) $X \sim$ Expo (λ) has $E(X) = 1/\lambda$, Var$(X) = 1/\lambda^2$, $M(t) = (1-t/\lambda)^{-1}$. Now let $S_n = X_1 + \ldots + X_n$.
Then $E(S_n) = n/\lambda$, Var $(S_n) = n/\lambda^2$, $S_n^* = \{S_n - (n/\lambda)\}/(n/\lambda^2)^{1/2}$.

$$M(t:S_n^*) = M\left(t:\frac{\lambda}{\sqrt{n}}S_n - \sqrt{n}\right) = e^{-\sqrt{n}t}M\left(\frac{\lambda t}{\sqrt{n}}:S_n\right) = e^{-\sqrt{n}t}\left[M\left(\frac{\lambda t}{\sqrt{n}}:X_i\right)\right]^n = e^{-\sqrt{n}t}\left(1-\frac{\lambda t}{\sqrt{n}\lambda}\right)^{-n}$$

$$\therefore \ \ln M(t:S_n^*) = -\sqrt{n}\,t - n\log(1-t/\sqrt{n}) = -\sqrt{n}\,t + n[t/\sqrt{n}+t^2/2n+O(1/n)] = t^2/2+O(1/\sqrt{n})$$

$$\lim_{n\to\infty}\log M(t:S_n^*) = t^2/2 \Rightarrow \lim_{n\to\infty} M(t:S_n^*) = e^{t^2/2} = \text{m.g.f. of } N(0,1).$$

Conclusion. The limiting distribution of S_n^* is $N(0, 1)$ which proves the C.L.T.

Method of Approximations by C.L.T.

1. Construct $S_n^* = [S_n - E(S_n)]/\sqrt{\text{Var}(S_n)}$, where $S_n = X_1 + X_2 + \ldots + X_n$.

2. For large n, treat S_n^*, as $Z \sim N(0, 1)$, Notate $\Phi(t) =$ c.d.f. of $N(0, 1)$

Two Illustrations :

(i) Let $X \sim$ bin (n, p) Treat $X = X_1 + \ldots + X_n$ where $X_i \sim$ Ber (p) so that

$$S_n^* = (X-np)/\sqrt{npq}$$

$$\therefore \quad P\{X \le x\} = P\{S_n^* \le (x-np)/\sqrt{npq}\} \approx \Phi[(x-np)/\sqrt{npq}]$$

A slightly improved result follows by using *continuity correction* :

$$P\{X \le x\} \approx \Phi[(x+0.5-np/\sqrt{npq}).$$

(ii) Let $X \sim$ Pois (λ); Treat $X = X_1 + X_2 + \ldots + X_n$ where $X_j \sim$ Pois (λ/n).

$$S_n^* = (X-\lambda)/\sqrt{\lambda}$$

$$\therefore \quad P\{X \le x\} = P\{S_n^* \le (x-\lambda)/\sqrt{\lambda}\} \approx \Phi[(x-\lambda)/\sqrt{\lambda}]$$

Use Continuity correction to improve this approximation :

$$P\{X \le x\} \approx \Phi[(x+0.5-\lambda)/\sqrt{x}].$$

17-20. *The Lindeberg-Feller C.L.T.*

Let $X_1, X_2, \ldots, X_n$ be independent non-degenerate variates with distribution functions $F_1, F_2, \ldots, F_n$ means $\mu_1, \mu_2, \ldots, \mu_n$ and finite variates $\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2$; respectively. Let

$$B_n^2 = \text{Var}\,(S_n) = \text{Var}\,(X_1 + \ldots + X_n) = \sigma_1^2 + \sigma_2^2 + \ldots + \sigma_n^2.$$

If F_i are absolutely continuous with p.d.f. 'f_i' and if the Lindberg-Feller (L-F) condition

$$\lim_{n\to\infty} \frac{1}{B_n^2} \sum_{i=1}^{n} \int_{A_n} (x-\mu_i)^2 f_i(x)\,dx = 0, \qquad (A_n = |X-\mu_i| > \varepsilon B_n)$$

holds for all $\varepsilon > 0$, then

$$\lim_{n\to\infty} P\{S_n^* \le a\} = \lim_{n\to\infty} P\left\{\left[\frac{S_n - \mathsf{E}(S_n)}{B_n}\right] \le a\right\} = \Phi(a) = \text{c.d.f. of } N(0,1).$$

Proof is omitted.

17-21. *Lyapounov's C.L.T.*

Let $X_1, \ldots, X_n$ be *independent* variates with $\mathsf{E}(X_i) = \mu_i$ and $\text{Var}\,(X_i) = \sigma_i^2$ and $B_n^2 = \sigma_1^2 + \ldots + \sigma_n^2$. If a positive number δ can be found such that

$$\lim_{n\to\infty} B_n^{-(2+\delta)} \sum_{i=1}^{n} \mathsf{E}(|X_i - \mu_i|^{2+\delta}) \equiv \lim_{n\to\infty} \frac{\nu_{2+\delta}}{(B_n)^{2+\delta}} = 0 \qquad \text{[Lyapounov Condition]} \quad \ldots(1)$$

then $$\lim_{n\to\infty} P\{S_n^* \le a\} = \lim_{n\to\infty} P\left\{\left[\frac{S_n - \mathsf{E}(S_n)}{B_n}\right] \le a\right\} = \Phi(a) = \text{c.d.f. of } N(0,1).$$

Proof. We check that Lyapounov condition implies that Lindberg-Feller condition is satisfied. So consider L-F condition §17-20.

$$\frac{1}{B_n^2} \sum_{i=1}^{n} \int_{A_n} (x-\mu_i)^2 f_i(x)\,dx = \frac{1}{B_n^2} \sum_{i=1}^{n} \int_{A_n} \frac{|x-\mu_i|^{2+\delta}}{|x-\mu_i|^{\delta}} f_i(x)\,dx$$

$$\le \frac{(B_n)^{-2}}{(\varepsilon B_n)^{\delta}} \sum_{i=1}^{n} \int_{A_n} |x-\mu_i|^{2+\delta} f_i(x)\,dx \qquad \{\because |x-\mu_i| > \varepsilon B_n \text{ on } A_n\}$$

$$\le \frac{(B_n)^{-(2+\delta)}}{(\varepsilon)^{\delta}} \sum_{i=1}^{n} \int_{-\infty}^{\infty} |x-\mu_i|^{2+\delta} f_i(x)\,dx, \qquad \{\because A_n \subset R_{Xi}\}$$

$$= \frac{(B_n)^{-(2+\delta)}}{(\varepsilon)^{\delta}} \sum_{i=1}^{n} \mathsf{E}(|X-\mu_i|^{2+\delta}) \to 0, \text{ as } n\to\infty, \qquad \text{[by (1)]}$$

Hence Lindberg-Feller condition is satisfied and so C.L.T. holds.

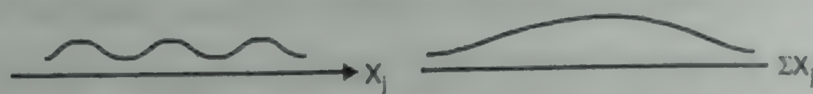
Note. In numerical problems, we may take $\delta = 1$. Thus for C.L.T to hold,

$$\lim [v_3 (B_n)^3] = 0, \text{ as } n \to \infty \quad ...(2)$$

Lyapounov Bound. Let $\lambda_n = v_{2+\delta} / B_n^{2+\delta}$. Then

$$| P(S_n^* \le t) - \Phi(t) | \le 0.8\, \lambda_n. \qquad [\delta = 1 \text{ suffices}] \quad ...(3)$$

This bound is too general to provide sharp results for specific cases : $\lambda_n \to 0$ is Lyapuonov condition.



Remarks. The essence of C.L.T is that repeated samples of size n, even from a non-normal population, generate an approximately normal distribution of either sums or means as indicated in the above figures. All samples $x_1, x_2, ..., x_n$ from a non-normal population X_j, which is far from being symmetric, generate an approximately normal population of Σ-sums X_j.

Normal approximation. Let $E(X) = \mu$ and Var $(X) = \sigma^2$. By the normal approximation to X, we mean the approximation $F_X(t) = \Phi[(t - \mu)/\sigma]$.

17-30. *Theorem*

Let $Z_1, Z_2, ..., Z_n$ be i.i.d variates with $E(Z_i) = \mu \ne 0$ and Var $(Z_i) = \sigma^2 \ne 0$. If $S_n = Z_1 + ... + Z_n$, then $P\{a < S_n < b\} \to 0$, as $n \to \infty$, where a and b are any two real numbers, with $b > a$.

Proof. Let $S_n^* = [S_n - E(S_n)]/\sigma(S_n) = (S_n - n\mu)/\sigma\sqrt{n}$, so that $S_n = n\mu + \sigma\sqrt{n}\, S_n^*$. Thus

$$\{a < S_n < b\} = \{a < n\mu + \sigma\sqrt{n}\, S_n^* < b\} = \{A_n < S_n^* < B_n\}, \quad ...(1)$$

where $$A_n = (a^{-1} - \mu)\sqrt{n}/\sigma, \quad B_n = (bn^{-1} - \mu)\sqrt{n}/\sigma.$$

Case 1. $\mu = 0$. Let $\varepsilon > 0$ be an arbitrary number; choose a positive number h so that

$$H = \int_{-h}^{h} \frac{e^{-u^2/2}}{\sqrt{2\pi}}\, du < \frac{\varepsilon}{2} \quad ...(2)$$

Now choose a positive integer N_1 so that $A_n > -h$, $B_n < h$ whenever $n > N_1$. ...(3)

Choose a positive integer N_2 so that

$$| P(-h < S_n^* < h) - H | < \tfrac{1}{2}\varepsilon, \text{ for } n > N_2 \quad ...(4)$$

The C.L.T. assures us that we can find such a positive integer N_2. Now suppose $n > N$, where $N = \max\{N_1, N_2\}$. From (1) and (3) it follows that $(a < S_n < b) = (A_n < S_n^* < B_n) \subset \{-h < S_n^* < h\}$.

Thus $P(a < S_n < b) \le P(-h < S_n^* < h) < H + \tfrac{1}{2}\varepsilon < \tfrac{1}{2}\varepsilon + \tfrac{1}{2}\varepsilon = \varepsilon,$ by (2) if $n > N$

$\therefore \quad \lim\limits_{n \to \infty} P(a < S_n^* < b) = 0.$ [ε is arbitrarily small]

Case 2. Let $\mu > 0$. Let $\delta > 0$ be an arbitrary number. Choose a negative number $-h$ such that

$$H_1 \equiv \int_{-\infty}^{-h} \frac{e^{-u^2/2} du}{\sqrt{2\pi}} < \frac{\delta}{2}. \quad ...(5)$$

Since $\lim A_n \to -\infty$, $\lim B_n \to -\infty$, we can choose a positive integer m_1 such that

$$A_n < -h, \quad B_n < -h, \quad \text{for } n > m_1. \qquad \ldots(6)$$

Using C.L.T. we can choose a positive integer m_2 such that

$$| P(S_n^* < -h) - H_1 | < \tfrac{1}{2}\delta, \text{ for } n > m_2. \qquad \ldots(7)$$

From (1) and (6) we conclude that

$$\{a < S_n < b\} = \{A_n < S_n^* < B_n\} \subset \{S_n^* < -h\}$$

$$\therefore \quad P(a < S_n < b) \le P(S_n^* < -h) < H_1 + \tfrac{1}{2}\delta < \tfrac{1}{2}\delta + \tfrac{1}{2}\delta = \delta \qquad \text{[by (7) and by (5)]}$$

This amounts to : $\lim P(a < S_n < b) = 0$, as $n \to \infty$ when $\mu > 0$ [δ is arbitrary small]. The case $\mu < 0$ is similar.

Cor. $P\{a < S_k < b, \ \forall\ k\} = 0.$ $\qquad \left[\text{i.e. } P\left\{\bigcap_{k=1}^{\infty}(a < S_k < b)\right\} = 0.\right] \quad \ldots(8)$

Proof. If n is any particular positive integer, then surely

$$\{a < S_k < b, \ \forall\ k\} \subset \{a < S_n < b\} \Rightarrow P\{a < S_k < b, \forall\ k\} \le P\{a < S_n < b\}. \quad \ldots(9)$$

The relation (9) is true for each positive integer n. The R.H.S. of (9) approaches 0 as $n \to \infty$ (Theorem 17-30) while the L.H.S. is independent of n. Thus, the L.H.S. is a non-negative number smaller than every positive number and hence it must be zero. So (8) holds

17-31. *Summary of Selected Results for C.L.T.*

Let $\{X_n, n > 1\}$ be a sequence of *independent* variates defined on some probability space. Denote :

$$S_n = X_1 + \ldots + X_n, \ \mu_k = \mathsf{E}(X_k), \ \sigma_k^2 = \text{Var}(X_k), \ B_n^2 = \text{Var}(S_n) = \sigma_1^2 + \ldots + \sigma_n^2 = \sigma^2(S_n).$$

The sequence $\langle X_n \rangle$ satisfies C.L.T. if

$$\lim_{n\to\infty} P\{S_n^* \le x\} = \Phi(x) = [\text{c.d.f. of } N(0,1)]. \qquad S_n^* = \frac{S_n - \mathsf{E}(S_n)}{\sigma(S_n)}.$$

Let F_n denote the c.d.f. of X_n and assume that $\mathsf{E}(X_n) = 0$ for all $n \ge 1$. We record the following three conditions.

(L). *Lindberg Condition* : $\lim\limits_{n\to\infty} \dfrac{1}{B_n^2} \sum\limits_{k=1}^{n} \int_A u^2 d\,F_k(u) = 0,$ for each $\varepsilon > 0$. $A = \{\ |u| \ge \varepsilon\, B_n\}$.

(*F*). *Feller Condition* : $\lim\limits_{n\to\infty} \max\limits_{1\le k\le n} \left(\dfrac{\sigma_k^2}{B_n^2}\right) = 0.$

(UAN). UAN Condition : $\lim\limits_{n\to\infty} \max\limits_{1\le k\le n} P\{|X_k / B_n| \ge \varepsilon\} = 0.$

[U.A.N. = Uniform asymptotic negligibility]

Two fundamental Theorems are :

Lindberg Theorem : $(L) \Rightarrow$ (C.L.T.)

Lindberg-Feller Theorem : $\{(L) \Leftrightarrow$ (C.L.T.) & $(F)\}$ or $\{(L) \Leftrightarrow$ (C.L.T.) & (U.A.N.)$\}$

17-32. Some Theoretical Problems Worked-out

***Example* 1.** Let $X_1, X_2, ..., X_k$ be independent bin (n_i, p_i), $1 \le i \le k$ variates. If n_i and k are large enough, show that

$$\sum_{i=1}^{n} |X_i - n_i p_i| \sim N\left[\sqrt{\frac{2}{\pi}} \sum_{i=1}^{k} \sqrt{n_i p_i q_i}, \left(1-\frac{2}{\pi}\right) \sum_{i=1}^{k} n_i p_i q_i\right].$$

Solution. If X is $N(0, \sigma^2)$, then $Y = |X|$, called folded normal variate, has

$$E(Y) = \sigma/\sqrt{2/\pi}, \qquad \text{Var}(Y) = \sigma^2.[1-(2/\pi)]. \qquad ...(1)$$

Now, let $Y_i = |X_i - n_i p_i|$. For large n_i

$$(X_i - n_i p_i) \sim N(0, \sigma_i^2); \; \sigma_i^2 = n_i p_i q_i \qquad ...(2)$$

$$S_k = Y_1 + Y_2 + ... + Y_k; \quad E(S_k) = (\Sigma\sigma_i)\sqrt{2/\pi}, \quad \text{Var}(S_k) = [1-2/\pi](\Sigma\sigma_i^2)$$

where we have used (1), the folded normal values.

By C.L.T., $[S_k - E(S_k)]/\sigma_{S_k} \sim N(0, 1)$

This means, $\qquad S_k \sim N[E(S_k), \text{Var}(S_k)] \qquad ...(3)$

This is, in essence, the result sought.

***Example* 2.** Let $X \sim$ Cauchy $(0, b)$ and $X_1, X_2, ..., X_n$ be a random sample of X. Let $Y_n = (X_1 + X_2 + ... + X_n)/n$. Find the p.d.f. of Y_n. Does the C.L.T. hold for Y_n ?

Solution. (*i*) We know that the p.d.f. and Ch. Function of X are

$$f(x) = \frac{b/\pi}{b^2 + x^2}, \; -\infty < x < \infty, \qquad \phi_X(t) = e^{-b|t|} \qquad ...(1)$$

$$\therefore \qquad \phi(t = Y_n) = \phi(t/n : X_1 + ... + X_n) = [\phi(t/n : X_1)]^n = e^{-nb|t/n|} = e^{-b|t|}$$

This shows that Y_n has the same p.d.f. as X, i.e. [1*a*]

(ii) Since $\phi(t, Y_n)$ is independent of n and so is its p.d.f., it follows that Y_n does not tend to a normal r.v. as $n \to \infty$. Consequently, the C.L.T. does not hold in the Cauchy distribution.

Note. Since m.g.f. for Chy $(0, b)$ does not exist, we are to use Ch. Function.

***Example* 3.** Examine if C.L.T. holds for the following sequences of indep. variates.

(a) $\qquad P\{X_k = \pm 2^k\} = 2^{-(2k+1)}, P\{X_k = 0\} = 1 - 2^{-2k}$

(*b*) $\qquad P\{X_k = \pm k^{\alpha}\} = \frac{1}{2} k^{-2\alpha}, P\{X_k = 0\} = 1 - k^{1-2\alpha}, \; \alpha < \frac{1}{2}.$

Solution. For the non i.i.d. variates, for C.L.T. to hold, we must have lim $\lambda_n = \lim(v_3 / B_n^2) = 0$, as $n \to \infty$. [Lyapounov's Condition] ...(1)

(a)
$$E(X_k) = 2^{-(2k+1)} 2^k + 2^{-(2k+1)}(-2^k) + (1 - 2^{-2k})0 = 0$$

$$E(X_k^2) = 2^{-(2k+1)} 2^{2k} + 2^{-(2k+1)} (2^{2k}) = 1, \text{ so Var}(X_k) = 1.$$

$$E(|X_k|^3) = 2^{-(2k+1)} 2^{3k} + 2^{-(2k+1)} 2^{3k} = 2^k$$

$$(B_n^2) = \sigma_1^2 + \sigma_2^2 + \ldots + \sigma_n^2 = n,\ \ v_3 = \Sigma E(|X_k|^3) = 2 + 2^2 + \ldots + 2^n = 2(2^n - 1).$$

$$\lim_{n\to\infty} \frac{v_3}{(B_n)^3} = \lim_{n\to\infty} \frac{2(2^n-1)}{n^{3/2}} = \lim_{n\to\infty} \frac{2(2^n \ln 2)}{(3/2)n^{1/2}} = \lim_{n\to\infty} \frac{2(2^n \ln^2 2)}{(3/4)\,n^{-1/2}} \to \infty.$$

where we have used L'Hopital Rule to evaluate indeterminates.

Since condition (1) is violated. C.L.T. does not hold for this sequence.

(b)
$$\begin{aligned} E(X_k) &= \tfrac{1}{2}k^{-2\alpha}.k^{\alpha} + \tfrac{1}{2}k^{-2\alpha}(-k^{\alpha}) + (1-k^{1-2\alpha}).0 = 0 \\ E(X_k^2) &= \tfrac{1}{2}k^{-2\alpha}.k^{2\alpha} + \tfrac{1}{2}k^{-2\alpha}(k^{2\alpha}) = 1, \text{ so Var}(X_k) = 1. \\ E(|X_k|^3) &= \tfrac{1}{2}k^{-2\alpha}.k^{3\alpha} + \tfrac{1}{2}k^{-2\alpha}.k^{3\alpha} = k^{\alpha} \\ (B_n)^2 &= \sigma_1^2 + \sigma_2^2 + \ldots + \sigma_n^2 = n,\ \ v_2 = \Sigma E(|X_k|^3) = 1^{\alpha} + 2^{\alpha} + \ldots + n^{\alpha}. \end{aligned}$$

Since each term in the last sum $\le n^{\alpha}$, we have $v_2 \le n(n^{\alpha}) = (n)^{\alpha+1}$.

$$\therefore \quad \lim_{n\to\infty} \frac{v_2}{(B_n)^3} \le \lim_{n\to\infty} \frac{n^{\alpha+1}}{n^{3/2}} = \lim_{n\to\infty} \frac{1}{n^{1/2-\alpha}} \to 0 \quad (\because \alpha < \tfrac{1}{2})$$

Since Lyapounov's condition (1) is satisfied, the C.L.T. holds for the sequence in (b).

***Example* 4.** Does the WLLN follow from Lindeberg Levy C.L.T.

Solution. The WLLN for the i.i.d. sequence $\langle X_n \rangle$ states that

$$\lim_{n\to\infty} P\{S_n / n) - \mu| < c\} = 1, \ (c > 0). \qquad \ldots(1)$$

Since c is fixed, for any positive k (constant), $k\sigma\sqrt{n} < cn$ for all sufficiently large values of n (because any large multiple of $\sqrt{n}$ is negligible in comparision with the small multiple of n). This provides for the following sub-event relation.

$$\{|S_n - n\mu| \le k\sigma\sqrt{n}\} \subseteq \{|S_n - n\mu| < nc\}$$

$$\therefore \quad P\{|S_n^*| < k\} = P\left\{\left|\frac{S_n - n\mu}{\sigma\sqrt{n}}\right| < k\right\} \le P\left\{\left|\frac{S_n - n\mu}{n}\right| < c\right\}, \text{ for large } n \qquad \ldots(2)$$

As $n \to \infty$, the left member converges to $2\psi(k)$, by C.L.T. Given any $\delta > 0$, we can first choose k so large that the value of $2\psi(k)$ exceeds $1 - \delta$; then choose n so large that relation (2) holds. It follows that

$$P\{|S_n/n) - \mu| < c\} > 1 - \delta$$

for all sufficiently large n. This is equivalent to (1). Thus, the WLLN is a corollay to C.L.T.

Note. $\displaystyle \lim_{n\to\infty} P\left\{\left|\frac{S_n - E(S_n)}{n}\right| < c\right\} = \lim_{n\to\infty} P\left\{\left|\frac{S_n - E(S_n)}{\sigma(S_n)}\right| < \frac{c}{\sigma/\sqrt{n}}\right\} = \lim_{n\to\infty} P\left\{|S_n^*| < \frac{c\sqrt{n}}{\sigma}\right\} = \int_{-\infty}^{\infty} f(t)\,dt = 1.$

***Example* 5.** A sequence of independent variates $\langle X_k \rangle$ is defined by

$$P(X_k = 0) = 1 - k^{-2c},\ \ P(X_k = k^c) = \tfrac{1}{2}k^{-2c},\ \ P(X_k = -k^c) = \tfrac{1}{2}k^{-2c}.$$

Show that the Lyapounov condition is satisfied only if $c < 1/2$, and the variable S_n is asymptotically $N(0, n)$.

Solution. $E(X_k) = 0(1-k^{-2c}) + k^c.\left(\frac{1}{2}k^{-2c}\right) + (-k^c)\left(\frac{1}{2}k^{-2c}\right) = 0$; $E(X_k^2) = k^{2c}.\left[\frac{1}{2}k^{-2c} + \frac{1}{2}k^{-2c}\right] = 1$

Thus, $\quad \text{Var}(X_k) = E(X_k^2) - E^2(X_k) = 1; \; E(S_n) = 0, B_n^2 = \text{Var}(S_n) = \text{Var}(\Sigma X_k) = n.$

Now $\quad E(|X_k - \mu_k|^3) = E(|X_k|^3) = k^{3c}.\left(\frac{1}{2}k^{-2c} + \frac{1}{2}k^{-2c}\right) = k^c.$

$$v_3 = \Sigma(E|X_k|^3) = 1 + 2^c + 3^c + \ldots + n^c \le n(n^c). \qquad [\text{if } c \ge 0]$$

Thus, $\quad \lambda_n = (v_3 / B_n^3) \le n^{c+1} / n^{3/2} = n^{c-(1/2)}.$

If $c < 1/2$, $\lambda_n \to 0$ as $n \to \infty$ ($\lambda \not\to 0$), so Lyapounov condition is satisfied.
If $c > 1/2$, λ_n is unbounded. For $c = 1/2$, we note that

$$\lambda_n = \left(1 + \sqrt{2} + \sqrt{3} + \ldots + \sqrt{n}\right) / n^{3/2} = (1/n)\,\Sigma\left(\sqrt{r/n}\right) = \int_0^1 \sqrt{x}\,dx = 2/3 \ne 0.$$

Thus, Lyapounov condition is not satisfied. For large n, $S_n \sim N(0, n)$. This we may verify:

$$M(t : X_n) = E(e^{tX_n}) = (1 - n^{-2c}) + \tfrac{1}{2}n^{-2c}\,(e^{tn^c} + e^{-tn^c}) = 1 - n^{-2c} + n^{-2c}\cosh(n^c t)$$

$$M(t : S_n^*) = M(t : S_n / \sqrt{n}) = M(t/\sqrt{n} : S_n) = [M(t/\sqrt{n} : X_n)]^n$$

$$= [1 - n^{-2c} + n^{-2c}\cosh(n^c t/\sqrt{n})]^n = [1 + t^2/2n + t^4/n^{(2-2c)}4! + \ldots]^n$$

$$\lim_{n\to\infty} M(t : S_n^*) = \lim_{n\to\infty}\left[1 + \frac{t^2}{2n} + O\left(\frac{1}{n}\right)\right]^n = e^{t^2/2} = [\text{m.g.f. of } N(0, 1)], \text{ if } c < 1/2. \text{ [Euler's limit]}$$

Thus, $\quad S_n^* \sim N(0,1)$, i.e. $(S_n / n^{1/2}) \sim N(0,1)$, i.e. $S_n \sim N(0, n)$.

Problems with Solutions Provided at the End of the Text

1*. Let $X \sim \text{Pois}(3n)$ and define $Y = (X - 3n)/\sqrt{3n}$. Using C.L.T. show that the limiting distribution of Y is $N(0, 1)$.

2*. If $X_1, \ldots, X_n$ are i.i.d variates, with p.m.f. $P(X_i = \pm 1) = 1/2$, show that C.L.T. holds for this sequence.

3*. If $X_1, X_2, \ldots, X_n$ are i.i.d, with $E(X_i) = 0$, $\text{Var}(X_i) = \sigma^2$, that $0 < \sigma^2 < \infty$, show that if $\varepsilon > 0$ be real $(\overline{X} = \Sigma X_i / n)$.

$$P\{\overline{X} \ge \varepsilon\} \approx \frac{\sigma}{\sigma\sqrt{n}}\,\frac{1}{2\pi}\exp\left\{\frac{-n\varepsilon^2}{2\sigma^2}\right\} \text{ for large } n. \qquad \ldots(1)$$

4*. Prove that if $\{X_n, n \ge 1\}$ are i.i.d. variates, then C.L.T. is a stronger result than WLLN.

5*. Let $X_1, X_2, \ldots, X_n$ be i.i.d random variables with $E(X_1) = \mu$, $\text{Var}(X_i) = \sigma^2 < \infty$, $E(X_1 - \mu)^4 = 1 + \sigma^4$. Now state C.L.T. and evaluate

$$\lim_{n\to\infty} P\left\{\sigma^2 - \frac{1}{\sqrt{n}} \le \sum_{i=1}^{n}\frac{(X_i - \mu)^2}{n} \le \sigma^2 + \frac{1}{\sqrt{n}}\right\}. \qquad \ldots(1)$$

6*. $\overline{X}_n$ is the mean of a random sample of size n from $U(0, 1)$ population. Let $T_n = \sqrt{n}\left(\overline{X}_n - \frac{1}{2}\right)$. Show that $\mathcal{F}(t) = \lim P\{T_n \le t\}$ as $n \to \infty$, exists and find it.

7*. Let $\langle X_n \rangle$ be a sequence of non-negative i.i.d variates with $E(\ln X_1)^2 < \infty$. Let $Y_n = (X_1, X_2, \ldots, X_n)^{1/n}$. Show that a constant $k > 0$ can be so chosen that the variate $(kY_n)^{\sqrt{n}}$ has non-degenerate limit distribution function and find it.

8*. Give an example of a distribution which satsifies Lindberg condition but does not satisfy Lyapunov condition.

9*. Construct a sequence that obeys C.L.T. but for which the three conditions [(L), (F) & (UAN)] do not hold.

Exercise 17(a)

1. Let $X_1, X_2, \ldots, X_n$ be i.i.d. variates with p.d.f. $f(x) = |x|, -1 \le x \le 1, f(x) = 0$, otherwise. Show that C.L.T. holds for the sequence $\{X_n\}$.
2. Let $X_1, X_2, \ldots, X_n$ be a sequence of i.i.d. random variables. Show that C.L.T. holds for the following p.d.f.'s of sequence $\{X_n\}$

   (a) $P\{X = \pm 1\} = 1/4, P(X = 0) = 1/2$. (b) $f(x) = 1/2\, e^{-|x|}, -\infty < x < \infty$

   (c) $f(x) = 1/2, |x| \le 1, f(x) = 0$, otherwise.
3. Check whether the C.L.T. holds for the following sequences $\{X_k\}$ of independent variates :

   (a) $P\{X_k = \pm 2^k\} = 1/2$. [No]. (b) $P\{X_k = \pm k^a\} = 1/2$. [Yes]

   (c) $P\{X_k = \pm k\} = \frac{1}{2} k^{-1/2}$, $P(X_k = 0) = 1 - k^{-1/2}$ [Yes, LLN. No]

   (d) $P\{X_k = \pm 1\} = \frac{1}{2}(1 - 2^{-k}), P(X_k = \pm 2^k) = 2^{-k-1}$. [Yes];

   (e) $P\{X_k = \pm\sqrt{2k-1}\} = \frac{1}{2}$

   (f) $P(X_k = 1\} = p_k, P(X_k = 0) = 1 - p_k, 0 < p_k < 1$.

   (g) $P(X_1 = \pm 1) = \frac{1}{2}$ and for $k \ge 2$, $c > 1$.

   (h) $P\{X_k = \pm 1\} = \frac{1}{2} c, P\{X_k = \pm k\} = \dfrac{1}{2k^2}(1 - c^{-1}),\ P\{X_k = 0\} = 1 - (1/c) - k^{-2}(1 - 1/c)$.

   (i) $P\{X = \pm 2^k\} = 2^{-(2k+1)}, P(\{X_k = 0\} = 1 - 2^{-2k}$ [C.L.T. No; LLN. Yes]
4. Let $X_1, X_2, \ldots, X_n$ be an infinite sequence of independent variates with p.m.f..
   $P\{X_n = \pm (\ln n)^k\} = t^{-1}, P(X_n = 0) = 1 - 2t^{-1}$, where $t = (n + \alpha)\,[\ln (n + \alpha)]^\rho$
   ρ being a sufficiently large constant. Further ρ and k satisfy the inequality $2k + 1 > \rho$. Show that that the C.L.T. holds for $\rho < 1$. What happens if $\rho \ge 1$.
5. Let $Y_1, \ldots, Y_n$ be i.i.d variates with $P\{Y_1 = -1\} = P\{Y_1 = 1\} = 1/2$. Let $X_k = \sqrt{5}\, Y_k / 4^k$ and set $S_n = X_1 + \ldots + X_n$. Prove that $P\{|S_n| \le 1/2\} = 0$ for every $n \ge 1$. [This shows convergence to Φ is impossible].

6. Standard variates X and Y are independent with a common distribution P_X. If $aX + bY$ has the same distribution P_X, whenever $a^2 + b^2 = 1$, show that P_X corresponds to the normal density $\varphi(x)$.

7. Let Z_n be the normalized sum of n i.i.d lattice variable with span h and variance σ^2. Then for each x specified by $[(a - n\mu)/\sigma\sqrt{n}] + k[h/\sigma\sqrt{n}]$; $k = 0, \pm 1, \pm 2, \ldots$ show that

$$\lim_{n\to\infty} \frac{\sigma\sqrt{n}}{h} P\{Z_n = x\} = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}.$$

8. Show that C.L.T. need not hold for random sums of random variables.

17-40. *Cramer's Theorem*

Sequences $\langle X_n \rangle$ and $\langle Y_n \rangle$ of r.v.'s are such that $X_n \xrightarrow{L} X$, and $Y_n \xrightarrow{P} c, (c \neq 0)$, then $(X_n / Y_n) \xrightarrow{L} (X / c)$.

We ommit the proof but give some illustrations :

Example 1. Let $X_1, X_2, \ldots, X_{2n}$ be i.i.d. $N(0, 1)$ variates. Find the limiting distribution of $Z_n = U_n/V_n$ where

$$U_n = \left(\frac{X_1}{X_2} + \frac{X_3}{X_4} + \ldots + \frac{X_{2n-1}}{X_{2n}}\right). \quad V_n = X_1^2 + X_2^2 + \ldots + X_n^2.$$

Solution. Let $Y_j = X_{2j-1} / X_{2j}$; then Y_j is standard Chy; $U_n = \Sigma Y_j$ and $U_n/n = \Sigma(Y_j / n)$ is standard Chy. [It follows by Ch. Function]. We conclude that $(U_n / n) \xrightarrow{L} \text{Chy}(1,0)$.

Further $X_j^2 \sim \chi^2_{(1)}$; so $E(X_j^2) = 1$. We infer by Khintchin's WLLN that

$$(V_n/n)(\Sigma X_j^2)/n \xrightarrow{P} 1 (= E(X_1^2)).$$

$$\therefore \quad Z_n = \frac{U_n}{V_n} = \left(\frac{U_n / n}{V_n / n}\right) \xrightarrow{L} Z \sim \text{Chy}(1, 0), \text{ by Cramer's Theorem.}$$

Example 2. Let $X_1, X_2, \ldots, X_n$ be i.i.d., $N(0, 1)$ variates. Find the limiting distribution of

$$\sqrt{n}\,[(X_1 + X_2 + \ldots + X_n)] / (X_1^2 + X_2^2 + \ldots + X_n^2).$$

Solution. Let $U_n = (X_1 + X_2 + \ldots + X_n) / \sqrt{n}$; and $V_n = (X_1^2 + X_2^2 + \ldots + X_n^2) / n$.

Since $X_1^2 \sim \chi^2_{(1)}$, so $E(X_1^2) = 1$. Hence by Khintchin's WLLN, applied to sequence $\langle X_n^2 \rangle$, $V_n \xrightarrow{P} 1 [= E(X_1^2)]$.

$$M(t : U_n) = [M(t/\sqrt{n} : X_1)]^n = e^{+\frac{1}{2}t^2} \Rightarrow U_n \xrightarrow{L} Z \text{ where } Z \sim N(0, 1).$$

Hence by Cramer Theorem, $(U_n / V_n) \xrightarrow{L} Z$, where $Z \sim N(0, 1)$.

***Example* 3.** Let $X_1, X_2, \ldots, X_n$ be i.i.d. standardized variates with $E(X_i^4) < \infty$. Find the limiting distribution of $Z_n = \sqrt{n}\,[X_1 X_2 + X_3 X_4 + \ldots + X_{2n-1} X_{2n}]/(X_1^2 + X_2^2 + \ldots + X_n^2)$.

Solution.

Let $Y_j = X_{2j-1} X_{2j}$; Then $E(Y_j) = E(X_{2j-1})E(X_{2j}) = 0$ and $E(Y_j^2) = E(X_{2j-1})^2 E(X_{2j})^2 = 1$.

Thus Y_j is standard r.v. The Seq. $\{Y_j\}$ is an i.i.d. sequence. If $S_n = Y_1 + \ldots + Y_n$, then by CLT,

$$S_n^* = [(S_n - 0)/\sqrt{n}] \xrightarrow{L} Z, \text{ where } Z \text{ is } N(0, 1)$$

Consider the sequence $\langle X_j^2 \rangle$; $E(X_j^2) = 1$. Apply Khintchin's WLLN,

$$V_n = (X_1^2 + X_2^2 + \ldots + X_n^2)/n \xrightarrow{P} E(X_1^2) = 1.$$

Hence by Cramer's Theorem, $Z_n = \left(S_n^* / V_n\right) \xrightarrow{L} Z$, where $Z \sim N(0, 1)$.

***Exercise* 1.** Let $X_1, X_2, \ldots,$ be i.i.d. variates with $E(X_1) = a$ and $\text{Var}(X_1) = \sigma^2$ and

$$\bar{X}_n = \sum_{j=1}^{n} X_j / n.$$

Let $Y_1, Y_2, \ldots$ be i.i.d. variates with $E(Y_1) = b\,(\neq 0)$ and $\text{Var}(Y_1) = \tau^2$ and $\bar{Y}_n = \Sigma Y_i / n$.

Find the limiting distribution of $Z_n = \sqrt{n}\,(\bar{X}_n - a)/\bar{Y}_n$.

***Exercise* 2.** Let $X \sim$ Pois (λ) and $Y \sim$ Pois (μ) be independent. Prove or disprove :

$$\{[(X - Y) - (\lambda - \mu)]/(X + Y)^{1/2}] \xrightarrow{L} N(0, 1) \text{ as } \lambda \to \infty, \mu \to \infty.$$

17.50. *Continuity Correction*

When approximating a *discrete* probability law by continuous probability law, it is useful to adopt the following rule :

"Add 1/2 to the upper bound and subtract 1/2 from the lower bound". Thus, when X is discrete a more accurate evaluation of $P(a \le X \le b)$ is provided by $P\left\{a - \tfrac{1}{2} \le X \le b + \tfrac{1}{2}\right\}$.

Comments. Apply continuity corection only to *closed* intervals. For open or semi-open intervals first convert then to closed intervals and then apply continuity correction. Thus

$$P(a \le X \le b) = P(a - .5 \le Y \le b + .5);\quad P(X > a) = P(X \ge a + 1) \approx P(Y \ge a + 1 - .5)$$

$$P(a < X \le b) = P(a + 1 \le X \le b) \approx P(a + 1 - .5 \le Y \le b + 0.5)$$

$$P(a < X < b) = P(a + 1 \le X \le b + 1) \approx P(a + 1 - .5 \le Y \le b + 1 + 0.5)$$

Continuity correction is applied only to discrete variates, not to those who are already continuous.

17-51. *Normal Approximation to Binomial Distribution*

De-Moivre-Laplace Limit Theorem : *Normal approximation to Binomial Distn*

Let X_n bin (n, p) and Set $Z_n = (X_n - np)/\sqrt{npq}$. Then

$$P\{a \le Z_n \le b\} = P\{np + a\sqrt{npq} \le X_n \le np + b\sqrt{npq}\} \to \Phi(b) - \Phi(a), \text{ as } n \to \infty.$$

Proof. Show by MGF method that Z_n follows $N(0, 1)$ distribution. Then the

$$P\{a \le Z_n \le b\} = P\{Z_n \le b\} - P\{Z_n \le a\} = \Phi(b) - \Phi(a), \quad \text{as } n \to \infty.$$

17-52. *Classical Approach : Normal Approximation to the Poission Distribution*

If $X \sim$ Pois (λ) and if λ and X are large, then $U = (X - \lambda)/\sqrt{\lambda}$ is bounded. Also

$$u_{r+1} - u_r = \frac{(r+1) - \lambda}{\sqrt{\lambda}} - \frac{(r - \lambda)}{\sqrt{\lambda}} = \frac{1}{\sqrt{\lambda}}. \qquad \text{...(1)}$$

Hence, for large λ, $u_{r+1} - u_r = 1/\sqrt{\lambda} \Rightarrow du = \lambda^{-1/2}$.

Taking logarithms of $f(x) = e^{-\lambda}\lambda^x / x!$, and using Stirling formula for x, we obtain

$$\ln f(x) = x \ln \lambda - \lambda - \ln(x!) = x \ln \lambda - \lambda - \ln\left(\sqrt{2\pi}\, e^{-x}.x^{x+\frac{1}{2}}\right)$$

$$= x(\ln \lambda + 1) - \lambda - (\tfrac{1}{2} + x)\ln x - \ln\sqrt{2\pi}.$$

Putting $x = \lambda(1 + u/\sqrt{\lambda})$, $u/\sqrt{\lambda}$ being very small, we get

$$\ln g(u) = \lambda(1 + u/\sqrt{\lambda})(\ln \lambda + 1) - \lambda - \left(\tfrac{1}{2} + \lambda + u\sqrt{\lambda}\right)\ln \lambda(1 + u/\sqrt{\lambda}) + \text{const.}$$

$$= u\sqrt{\lambda} - \left[\tfrac{1}{2} + \lambda + u\sqrt{\lambda}\right]\ln(1 + u/\sqrt{\lambda}) - \ln\sqrt{2\pi\lambda}$$

$$= u\sqrt{\lambda} - \left[\tfrac{1}{2} + \lambda + u\sqrt{\lambda}\right] \qquad [u/\sqrt{\lambda} - u^2/2\lambda) + \ldots] - \ln\sqrt{2\pi\lambda}$$

i.e. $$\ln[\sqrt{2\pi\lambda}\, g(u)] = -\tfrac{1}{2}u^2 + O(1/\sqrt{\lambda}) \qquad \text{...(2)}$$

When $\lambda \to \infty$, using (1) we get $dg(u) = [\sqrt{2\pi})^{-1} \exp(-u^2/2)]\, du, \quad -\infty < u < \infty.$

Thus, a Poisson distribution with a large mean may be approximated by a normal distribution.

17-53. *Worked-out Problems*

Example 1. Use the normal approximation to the binomial distribution to determine the prob. of getting (i) 6 heads in 16 flips (ii) 20 heads in 40 flips, of a fair coin. Compare the results with the exact solutions.

Solution. To find the approximations, we need **continuity correction**, which requires to replace a non-negative integer k by the interval $[k - \frac{1}{2}, k + \frac{1}{2}]$

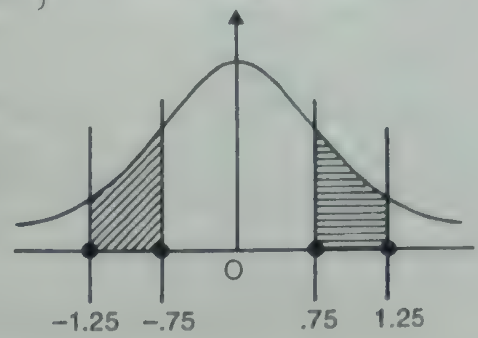
(*i*) Here $n = 16$, $x = 6$, p 1/2, $\mu = np = 8$; $\sigma^2 = npq = 4$.

$p_1 = P(X=6) = P\{5.5 \le X \le 6.5\} = P\left\{\frac{5.5-8}{2} \le \frac{X-\mu}{\sigma} \le \frac{6.5-8}{2}\right\} = P\{-1.25 \le Z \le -0.75\}$

$= P\{0.75 \le Z \le 1.25\}$, by symmetry

$= 0.3944 - 0.2734 = 0.1210$. [Table of Normal distribution]

The correct value from bin (n, p)-Table is 0.1222. Thus, the error of approx is -0.0012.



(ii) $n = 40$, $x = 20$, $p = 1/2$

$\mu = np = 20$, $\sigma^2 = npq = 10$, $\sigma = 3.1623$

$$p_2 = P\{X=20\} = P\{19.5 < X < 20.5\} = P\left\{\frac{19.5-20}{\sqrt{10}} < \frac{X-\mu}{\sigma} < \frac{20.5-20}{\sqrt{10}}\right\}$$

$$= P\{-0.158 < Z < 0.158\} = 2P\{0 < Z < 0.16\} \quad \text{by symmetry}$$

$$= 2(0.0636) = 0.1272. \text{ (Table-value)}$$

Actual value $P(X = 20) = \binom{40}{20}\left(\frac{1}{2}\right)^{40} = 0.1268.$ [using calculator]

***Example* 2.** If X is bin (n, p) and np is an integer, estimate the probability that X assumes the most probable value when n is large.

Solution. When n is large, the normal approximation of bionomial distribution, provided by C.L.T., is

$$P\{X=k\} = (\sqrt{2\pi npq})^{-1} \exp[-(k-np)^2/2npq]. \quad \text{...(1)}$$

Now the most probable value (*i.e.* modal value) of bin (n, p) is $[(n+1)p - 1]$ and if np is an integer, this value is just np. Hence, putting $k = np$ in Eq. (1) we get $P\{X = np\} = 1/\sqrt{2\pi npq}$.

***Example* 3.** Let $X_1, X_2, \ldots, X_n$ be a sequence of Bernoullis' variates with constant probability p of success in each trial. Using Normal approximation to Binomial distribution, find the smallest value of n for which $P\{|\bar{X}_n - p| \le \delta\}$ is at least γ, where δ and γ are fixed admissible constants. Compare this value with the value of n obtained using Chebyshev's inequality.

Solution. By Normal approximation to Binomial $Z = (\bar{X}_n - p)/\sqrt{pq/n} \sim N(0,1)$ for large n.

Here $\bar{X}_n = (1/n)\Sigma X_i$; $\mu = p$, $\sigma = \sqrt{pq/n}$;

$\therefore \quad P\{|\bar{X}_n - p| \le \delta\} = P\{|Z| \le \delta\sqrt{n}/\sqrt{pq}\} = 2\Phi[(\delta\sqrt{n}/\sqrt{pq})] - 1.$

Since $\sqrt{pq} \le (p+q)/2 = \frac{1}{2}$ (as $AM \ge G.M.$), so $\Phi(\delta\sqrt{n}/\sqrt{pq}) \le \Phi(2\delta\sqrt{n})$ because the c.d.f. Φ is an increasing function. Hence

$$p_0 = P\{|\bar{X}_n - p \le \delta\} \ge 2\Phi(2\delta\sqrt{n}) - 1.$$

Since p_0 is at least γ (*i.e.* min $p_0 \geq \gamma$), we get $2\Phi(2\delta\sqrt{n}) - 1 \geq \gamma \Rightarrow \Phi(2\delta\sqrt{n}) \geq \frac{1}{2}(\gamma + 1)$.

$$\therefore \quad 2\delta\sqrt{n} \geq z_{(\gamma+1)/2} \Rightarrow n \geq [z^2_{(\gamma+1)/2} / 4\delta^2] \Rightarrow \min n = (1/4\delta^2)\, z^2_{(\gamma+1)/2} \quad \ldots(1)$$

Now by Chebyshev's inequality : $P\{\bar{X} - p| \leq \delta\} \geq 1 - \sigma^2/\delta^2, [\sigma^2 = pq/n]$

i.e. $p_0 \geq 1 - (pq/n\delta^2) \geq 1 - (1/4n\delta^2)$. ($\because$ $pq < 1/4$).

Since p_0 is at least γ, so we must have

$$1 - (1/4n\delta^2) > \gamma \Rightarrow n > 1/4\delta^2(1-\gamma) \Rightarrow \min(n)\, 1/4(1-\gamma)\delta^2. \quad \ldots(2)$$

To compare (1) and (2), consider a concrete situation

$$p_0 = P\{|\bar{X}_n - p| < 0.01\} > 0.9.$$

Here $\delta = 0.01$, $\gamma = 0.9$, $z_{(\gamma+1)/2} = z_{.95}$. From $N(0, 1)$ areas.

$Z_{0.95}$ = 1.645 (critical value having 95% area on its left). Then by (1)

$$\min(n) = (0.25) \times 10000 \times (1.645)^2 = 2500 \times 2.706025 = 6765.$$

By Ghebyshev's inequality, by (2), min (n) = 10000/4 × 0.10 = 25000.

***Example* 4.** Let Y denote the sum of the items of a random sample of size 12 from a Die-distribution with p.d.f. $f(x) = 1/6$, $x = 1, 2, 3, 4, 5, 6,$; $f(x) = 0$, otherwise. Compute an approximation value of $P\{36 \leq Y \leq 48\}$.

Solution. Let $Y = X_1 + X_2 + \ldots + X_{12}$; $\mu = \mathrm{E}(Y) = 12\,\mathrm{E}(X)$; $\mathrm{Var}(Y) = 12\,\mathrm{Var}(X)$.

$$\mathrm{E}(X) = (1/6)(1 + 2 + \ldots + 6) = 7/2, \mathrm{E}(X^2) = (1/6)(1^2 + 2^2 + \ldots + 6^2) = 91/6.$$

$$\therefore \quad \sigma_X^2 = (91/6) - (49/4) = 35/12. \text{ So } \mathrm{E}(Y) = 42, \mathrm{Var}(Y) = 35.$$

Let $Z = (Y - \mu)/\sigma$, i.e. $Y = 42 + \sqrt{35}\, Z$ and $Z \sim N(0, 1)$. Now

$$p = P\{36 \leq Y \leq 48\} = P\{35.5 < Y < 48.5\} = P\{35.5 < 42 + \sqrt{35}\, Z < 48.5\} \text{ [Continuity Correction)}$$

$$= P\{-(6.5/\sqrt{35}) < Z < (6.5/\sqrt{35})\} = P\{|Z| \leq 1.1\} = 2\psi(1.1) = 2(0.3643) = 0.7386.$$

Problems with Solutions Provided at the End of the Text

1*. Two fair dice are tossed 600 times. Let X denote the number of times a total of 7 occurs. Using C.L.T., find $\{90 \leq X \leq 110\}$.

2*. A polling agency wishes to take a sample of voters in a given state large enough that probability is only 0.01 that they will find the proportion favouring a certain candidate to be less than 50% when infact it is 52%. How large a sample should be taken ?

3*. A coin is tossed 10 times. Find the probability of 3, 4 or 5 heads using (a) Binomial distribution (*b*) Normal approximation.

4*. A lady has 7 dark brown and 5 black (single) socks in a drawer. Each morning she wears a pair of socks at random, each evening she washes them and throws them back into the drawer. Find the probability that during one year (365 days) she will wear matching socks at least half of the time.

5*. At a distance of a yards from the explosion of a bomb, it is known that the probability of a pane of glass being smashed is 0.30. What is the probability that out of 100 panes of glass situated at a yards from the explosion, 40 or more shall be smashed ?

6*. Van of a truck can hold 25 containers (identical in shape and size). If the weight of a container is assumed to be a normal variate with mean 150 lb and S.D. 15 lb, what is the Prob. that 25 containers will overload the truck, if it can safely carry 4000 lbs.

7*. If X is $N(500, 100^2)$, find the 90th percentile and the quartiles for the Dist. of X.

Exercise 17(b)

1. (a) If X is bin (72, 1/3), find an approximate value of $P(22 \le Y \le 28)$.

   (b) If X is bin (100, 1/2), find an approximate value of $P(Y = 50)$

   (c) If X is bin (n, 0.55), find the smallest value of n so that, approx. $P\{(X/n) > \frac{1}{2}\} \ge 0.95$.

2. The round-off error to the second decimal place is $U(-1/20, 1/20)$. Find the probability that the absolute error in the sum of 1000 numbers is less than 2.

3. If 100 true coins are tossed, how would you obtain an approximation for the probability of getting (i) 55 heads (ii) 55 or more heads.

4. A gambler's daily income in rupees has U(–400, 500) distribution. Find the chance that he wins more than 5000 rupees in 60 days. What amount k (gain or loss) is such that at most once in 20 times on average, the gambler will have income less than k during 60 days ?

5. Let X denote the average of the values shown when n dice are rolled. Find the probability that the sum of 600 such averages lies in the interval $100 \pm 90n^{-1/2}$.

6. Using normal approximations, find the probability that 200 tosses of a true coin will result in (i) between 80 and 120 heads, both inclusive (ii) less than 90 heads (iii) less than 85 or more than 115 heads (iv) exactly 100 heads.

7. Show that the probability that the number of heads in 400 throws of a fair coin lies between 180 and 220 is approximately $2\Phi(2) - 1$.

8. A die is thrown 720 times. Find an approximate value of the probability of the following events : (1) 'Six' lies between 100 and 400 (ii) 'Six' comes for more than 130 times.

9. Apply the WLLN and CLT to the sequence $\langle X_n \rangle$ of independent variates indicating the eyes obtained at different throws of a die.

   If $n = 10000$, find the limits of S_n within which its random value may be chosen. Estimate $P\{S_n\}$ in these limits.

10. The number X of items of a certain kind demanded by customers follows Pois (9). What stock k of this item should a retailer keep, in order to have a probability 0.99 of meeting all demands made on him ? Use Normal approx to Poisson Law..

11. The number of people entering a store is Poissonly distributed with parameter 100 people per hour. How long should you wait in order to have a probability of 0.74, that more than 180 people have entered the store. ?

12. Let $X_1, \ldots, X_{15}$ be a random sample from $f(x) = 3x^2$, $0 \le x \le 1$; $f(x) = 0$, elsewhere. Find $P\{0.6 \le \overline{X}_{15} \le 0.8\}$ approximately.

13. Let $X_1, \ldots, X_{72}$ be a random sample from $f(x) = 1/x^2$, $0 < x < \infty$; $f(x) = 0$, elsewhere. Compute approximately the probability that more than 50 of the items of the random sample are less than 3.

14. A random walk on the x-axis is governed by the following rules. At each step, the change to the attained coordinate position is +1 with probability $\frac{1}{2}$ or $-\frac{1}{2}$ with probability $\frac{1}{2}$. Steps are taken independently of each other and the initial position is at the origin. Show that the probability that the walker is to the right of the origin after a large number n of steps is approximately $\Phi\left(\frac{1}{2}\sqrt{n}\right)$.

15. Obtain $M_X(t)$ when X is Pois (λ); find the limit as $\lambda \to \infty$ of m.g.f. of $(X-\lambda)/\sqrt{\lambda}$ and interpret the result in the context of the C.L.T. Also prove that

$$\lim_{n\to\infty} \sum_{k=\alpha}^{\beta} \frac{e^{-\lambda}\lambda^k}{k!} = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-u^2/2}\, du. \quad (\alpha = \lambda + a\sqrt{\lambda},\ \beta = \lambda + b\sqrt{\lambda})$$

Show that $2X$ is not Poisson Variate. Give a set of conditions under which $X + Y$ too is a Poisson Variate.

We all learn by experience, ... never lose sight of the alternative.

[Sherlock Holmes in the Adventures of the Black Peter]

Pain is life—the sharper, the more evidence of life. (Charles Lamb)

Uniform Distribution. Exponential Distribution 18

18-10. *Definition*

The random variable X, with a p.d.f. given by

$$f(x) = 1/(b-a),\ a \le x \le b,\ (b>a);\ f(x)=0,\ \text{otherwise} \qquad \ldots(1)$$

is said to be a *Uniform* (or *Rectangular*) variate on the interval (a, b) and $f(x)$ given by (1) is called *uniform* or *rectangular density function*. If X has the density function (1), it is expressed by writing $X \sim$ Unif (a, b). a and b are called *parameters* of the distribution. For uniform distribution, subintervals of (a, b) that have equal lengths, all have the same induced probability and hence all are equally likely. Thus the statement "choose a number at random in the interval (a, b) always mean that a number is the observed value of X where X is Unif (a, b). The probability that X takes on a value in the interval $(c, d) \subseteq (a, b)$ is given by

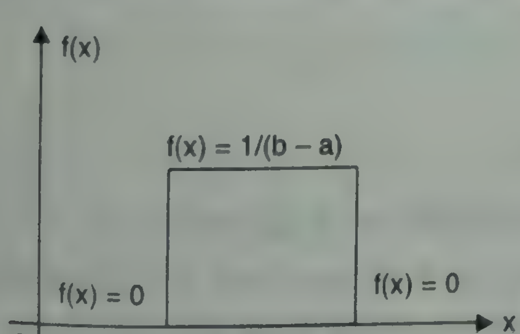
$$P = \int_c^d f(x)\,dx = \frac{d-c}{b-a},$$

P is called an *interval function.*

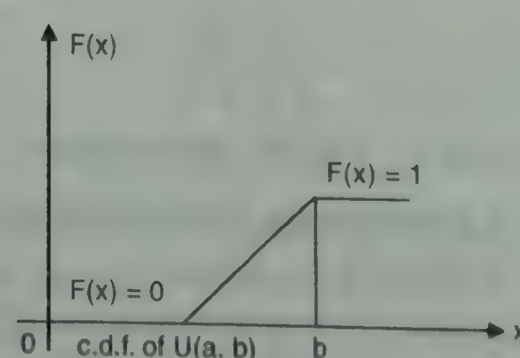
Distribution Function. If X is $U(a, b)$, the c.d.f. is

$$F_X(x) = 0,\ x < a\ ;\ F_X(x) = (x-a)/(b-a),\ a \le x \le b\ ;\ F(x) = 1,\quad x > b.$$

The graphic representation of f and F is as under :



p.d.f. of U(a, b)



c.d.f. of U(a, b)

***Example* 1.** Find the distribution function of the larger root of the equation $t^2 + 2t - X = 0$, where X is Unif (0, 2) distributed.

Solution. From given equation, the larger root is $Z = \sqrt{1+X} - 1$, $f_X(x) = \frac{1}{2}$, $0 < x < 2$.

Now $F_Z(z) = P\{Z \le z\} = P\{X \le z^2 + 2z\} = \frac{1}{2}\int_0^{z^2+2z} dx = \frac{1}{2}(z^2 + 2z)$

Since $0 \le x \le 2 \Rightarrow 0 \le \sqrt{1+x} - 1 \le \sqrt{3} - 1$, so the p.d.f. of Z is

$$F_z'(z) = f_z(z) = z+1, \quad 0 \le z \le \sqrt{3}-1; \quad f_Z(z) = 0, \text{ otherwise.}$$

***Example* 2.** Name three distributions for which $P(X \le \mu_x) = \frac{1}{2}$.

Solution. Trivially : bin $(2m+1,\ 1/2)$; $N(0, 1), U(0, 1)$

18-20. *Moments*

If μ_r' and μ_r denote rth order simple and central moments, then

$$\mu_r' = \mathsf{E}(X^r) = \int_a^b \frac{x^r\,dx}{(b-a)} = \frac{b^{r+1}-a^{r+1}}{(r+1)(b-a)}. \qquad \text{...(1)}$$

In particular, mean $\mu_1' = \frac{1}{2}(b+a)$

$$\mu_r = \mathsf{E}\{[X-\tfrac{1}{2}(a+b)]^r\} = \int_a^b \frac{[x-\frac{1}{2}(a+b)]^r}{b-a}\,dx = \frac{1}{b-a}\int_{-c}^{c} z^r\,dz, [z = x-\tfrac{1}{2}(a+b), c = \tfrac{1}{2}(b-a)]$$

Obviously, if $r = 2n+1$ (odd integer), the integer vanishes; hence

$$\mu_{2n+1} = 0, \quad n = 1, 2, 3, \ldots$$

If $r = 2n$ (even integer), the preceding integral reduces to

$$\mu_{2n} = \frac{2}{b-a}\int_0^c z^{2n}\,dz = \frac{(b-a)^{2n}}{(2n+1)\,2^{2n}} \qquad \text{...(2)}$$

In particular : $\quad$ Var $(X) = \mu_2 = (1/12)(b-a)^2, \mu_4 = (b-a)^4/80.$

The **absolute moments** follow easily :

$$\nu_r = \mathsf{E}\left(\left|X-\frac{b+a}{2}\right|^r\right) = \int_a^b \frac{|x-\frac{1}{2}(a+b)|^r}{b-a}\,dx = \frac{1}{b-a}\int_{-c}^{c}|z|^r\,dz = \frac{2}{b-a}\int_0^c z^r\,dz$$

$$= \frac{2}{b-a}\frac{c^{r+1}}{(r+1)} = \frac{(b-a)^r}{2^r(r+1)}. \qquad \text{...(3)}$$

Observe that ν_1 = M.D. about mean = $(b-a)/4$.

Pearson's Coefficients : $\beta_1 = \mu_3^2/\mu_2^3 = 0,\ \beta_2 = \mu_4/\mu_2^2 = 9/5;\ \gamma_1 = \sqrt{\beta_1} = 0, \gamma_2 = \beta_2 - 3 = -6/5.$

Since $\gamma_1 = 0$, the distribution $U(a, b)$ is symmetric (not skewed and $\gamma_2 < 0$ implies that it is Platycutric. Notice that the point (β_1, β_2) lies on the straight line $\beta_2 - \beta_1 - 3 = -6/5$.

18-21. *Moment Generating Function and Factorial M.G.F.*

$$M(t:X) = \mathsf{E}(e^{tX}) = \int_a^b e^{tx}\frac{dx}{b-a} = \frac{1}{b-a}\left(\frac{e^{bt}-e^{at}}{t}\right), \ t \ne 0. \qquad [\S\ 8\text{-}16\ (9)] \qquad \text{...(1)}$$

We expand Exponentials to get

$$M(t:X) = \frac{1}{b-a}\sum_{n=1}^{\infty}\left(\frac{b^n-a^n}{n!}\right)t^{n-1} = \frac{1}{(b-a)}\sum_{n=0}^{\infty}\left(\frac{b^{n+1}-a^{n+1}}{n+1}\right)\frac{t^n}{n!} \qquad \text{...(2)}$$

$$\therefore \quad \mu_n' = \text{Coeff. of } \frac{t^n}{n!} = \frac{1}{b-a}\left(\frac{b^{n+1} - a^{n+1}}{n+1}\right) \quad \ldots(3)$$

In particular, $\mu_1' = \frac{1}{2}(b+a)$, $\mu_2' = (b^3 - a^3)/3(b-a)$, $\sigma^2 = \mu_2' - \mu_1'^2 = \frac{1}{2}(b-a)^2$.

Characteristic Function : $M(it : X) = (e^{itb} - e^{ita})/it(b-a)$.

$$W(t) = \mathsf{E}[(1+t)^X] = \int_a^b \frac{(1+t)^x}{(b-a)}\,dx = \frac{(1+t)^b - (1+t)^a}{(b-a)\ln(1+t)}. \quad \text{[Factorial m.g.f.]}$$

18.22. *Mean Deviation about Mean, Median and Mode*

$$\text{M.D.} = \mathsf{E}(|X-\mu|) = \int_a^b \frac{|x - \frac{1}{2}(a+b)|}{(b-a)}\,dx = \frac{1}{b-a}\int_{-c}^{c} |z|\,dz$$

$$= \frac{2}{b-a}\int_0^c z\,dz = \frac{c^2}{b-a} = \frac{(b-a)}{4}. \quad [z = x - \tfrac{1}{2}(a+b),\ c = \tfrac{1}{2}(b-a)]$$

If m is the median, then

$$\frac{1}{2} = \int_a^m f(x)\,dx = \frac{1}{b-a}\int_a^m dx = \frac{m-a}{b-a} \Rightarrow m = \frac{1}{2}(a+b).$$

This result is otherwise obvious, since the distribution is symmetric about the mean $\frac{1}{2}(a+b)$.

Mode. Since in the rectangle $(b-a) \times (b-a)^{-1}$, each point in the interval (a, b) has the maximum probability, it follows that each point of (a, b) is a Mode. Some authors say that the mode does not exist.

Note. $\nu_r = \mathsf{E}(|X-\mu|)^r = (b-a)^r/(r+1)2^r$.

$$\mu_{2r+1} = 0, \qquad \mu_{2r} = (b-a)^{2r}/4^r(2r+1).$$

Example 1. Let $X \sim$ Unif $(-a, a)$. Determine a so that

(i) $P(-\leq X \leq 2) = 0.75$, (ii) $P(|X| < 1) = P(|X| > 2)$.

Solution. Here $f(x) = (1/2a)$. So that

(i) $0.75 = \int_{-1}^{2} f(x)\,dx = \dfrac{3}{2a} \Rightarrow a = 2.$

(ii) $\int_{-1}^{1} \dfrac{dx}{2a} = 1 - \int_{-2}^{2} \dfrac{dx}{2a} \Rightarrow \dfrac{1}{a} = 1 - \dfrac{2}{a} \Rightarrow a = 3.$

Example 2. Let X and Y be i.i.d. Unif (1, 2) variates. Compare $\mathsf{E}(X/Y)$ without using the density of X/Y.

Solution. Here $\mathsf{E}(X) = \int_1^2 x\,dx = \dfrac{3}{2}$, $\mathsf{E}(Y^{-1}) = \int_1^2 \dfrac{dy}{y} = \ln 2.$ $[f(x) = 1, \quad 1 \leq x \leq 2]$

$\therefore \quad \mathsf{E}(X/Y) = \mathsf{E}(X).\mathsf{E}(Y^{-1}) = (3/2)\ln 2.$ [X & Y are indep.]

18-30. *Probability-integral Transformation*

If the r.v. X has c.d.f. F_X, where F_X is continuous, then the r.v. $Y = F_X(x)$ is $U(0, 1)$.

Proof. Since $0 \leq F_X(x) \leq 1$, $\forall\, x$, we have $F_X(y) = 0$, for $y \leq 0$ and $F_X(y) = 1$, for $y \geq 1$. For $0 < y < 1$, define z to be the largest number satisfying $F_X(z) = y$. Then $F_X(x) \leq y$ iff $X \leq z$, and it follows that

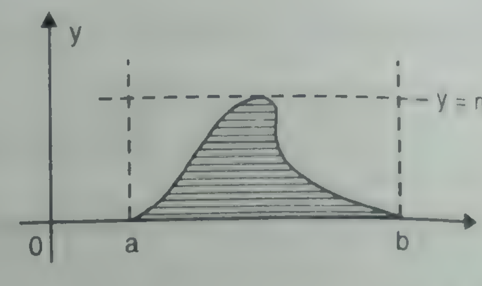
$$F_X(y) = P\{F_X(x) \le y\} = P\{X \le z\} = F_X(z) = y$$

which is the uniform distribution. [Another proof uses m.g.f. technique]

18-31. *Rejection Method for Continuous Variates*

(i) Suppose that $f(\cdot)$ is defined on a finite interval (a, b) and is bounded by m.

The graph of $f(x)$ is enclosed in the rectangle R : $a \le x \le b$, $0 \le y \le m$. We accept points (x, y), chosen at random in R, which lie in the shaded region and reject which lie outside.



For random points (x, y), we use distributions of X, as unif (a, b) and that of Y, independent of X, as unif $(0, m)$. Thus the joint p.d.f. is

$$f(x, y) = 1/m(b-a), \qquad a \le x \le b, \qquad 0 \le y \le m.$$

Write $\quad A = \{X \text{ accepted}\} = \{(x, y) : a \le x \le b, \quad 0 \le y \le f(x)\}.$

$$\therefore \quad P(A) = \int_a^b \left(\int_0^{f(x)} \frac{dy}{m(b-a)} \right) dx = \frac{1}{m(b-a)} \left(\int_a^b f(x)\, dx \right) = \frac{1}{m(b-a)}.$$

$$P\{X \le x_0, A\} = \int_a^{x_0} \left(\int_0^{f(x)} \frac{dy}{m(b-a)} \right) dx = \frac{1}{m(b-a)} \int_a^{x_0} f(x)\, dx = \frac{F(x_0)}{m(b-a)}$$

$$\therefore \quad P\{X \le x_0 \mid A\} = P\{X \le x_0, A\}/P(A) = F(x_0).$$

(ii) Suppose that $f(x)$ is defined on an infinite interval $]-\infty, \infty[$, $f(x)$ being bounded. Now we cannot choose a value from a distribution which is uniform over an infinite interval. Suppose however we can find another distribution, with p.d.f. : $g(\cdot)$ and a constant $c > 0$ such that

$$f(x) \le cg(x), \qquad [\forall\, x, g(x) > 0, f(x) > 0]$$

Now, a random value X is drawn from $g(.)$ and independently, a value U is drawn from unif $f(0, 1)$. If the realized pair x, u satisfy

$$u \le f(x)/cg(x) \qquad [\text{set } h(x) = f(x)/cg(x)]$$

we accept x ; the event A in this case. So

$$P(A) = \int_{-\infty}^{\infty} \left(\int_0^{h(x)} f(x, u)\, du \right) dx. \qquad [f(x, u) = f_1(x) . f_2(u) = g(x) . 1, \text{ by indep.}]$$

$$= \int_{-\infty}^{\infty} g(x) . \left(\int_0^h du \right) dx = \int_{-\infty}^{\infty} g(x) . h(x)\, dx = \frac{1}{c} \int_{-\infty}^{\infty} f(x)\, dx = \frac{1}{c}.$$

$$P\{X \le x_0, A\} = \int_{-\infty}^{x_0} \left(\int_0^{h(x)} f(x, u)\, du \right) dx = \frac{1}{c} \int_{-\infty}^{x_0} f(x)\, dx = \frac{F(x_0)}{c} \qquad [\text{As above}]$$

$$\therefore \quad P\{X \le x_0 \mid A\} = P\{X \le x_0, A\} / P(A) = F(x_0).$$

Note. $f(x) \le m \Rightarrow f(x) \le \{m . (b-a)\} \{1/(b-a)\} = cg(x)$

where $c = m(b - a)$ and $g(x) = 1/(b - a)$; which provides case (i) from (ii).

18-40. Formula for Sum of Uniform Random Variables

Let $X_1, X_2, \ldots, X_n$ be i.i.d. Unif (0, 1) variates and $S_n = X_1 + X_2 + \ldots + X_n$, $\overline{X}_n = S_n/n$.
Define : $C_x = x$ if $x > 0$ and $C_x = 0$, if $x \le 0$.
The p.d.f. of S_n is $g_n(x)$ and that of $\overline{X}_n$ is $h_n(y)$ where

$$g_n(x) = \frac{1}{\Gamma(n)} \sum_{r=0}^{n} (-1)^r \binom{n}{r} [C_{x-r}]^{n-1}. \quad r < x \le r+1 \qquad \ldots(1)$$

$$h_n(y) = \frac{n}{\Gamma(n)} \sum_{r=0}^{n} (-1)^r \binom{n}{r} [C_{ny-r}]^{n-1}. \quad \frac{r}{n} < y \le \frac{r+1}{n} \qquad \ldots(2)$$

Proof. If X is Unif (0, 1), then

$$M(-t : X) = \mathsf{E}(e^{-tX}) = \int_0^1 e^{-tx}\, dx = \frac{1-e^{-t}}{t}; \; M(-t : S_n) = \left(\frac{1-e^{-t}}{t}\right)^n. \qquad \ldots(i)$$

To show that $g_n(x)$ in the density of S_n we proceed as under :

$$\int_0^\infty e^{-tx}\, [C_{x-r}]^{n-1}\, dx = \int_r^\infty e^{-tx}(x-r)^{n-1}\, dx = e^{-rt} \int_0^\infty e^{-tz} . z^{n-1}\, dz = \frac{e^{-rt}\,\Gamma(n)}{t^n} . [x - r = z] \qquad \ldots(ii)$$

The m.g.f. corresponding to distribution $g_n(x)$ is

$$M(-t) = \int_0^\infty e^{-tx}\, g_n(x)\, dx = \sum_{r=0}^{n} (-1)^r \binom{n}{r} \int_0^\infty e^{-tx}\, [C_{x-r}]^{r-1} \frac{dx}{\Gamma(n)}$$

$$= \sum_{r=0}^{n} (-1)^r \binom{n}{r} \frac{e^{-rt}}{t^n} = (1-e^{-t})^n / t^n, \qquad \text{[by (ii)]} \qquad \ldots(iii)$$

Since (iii) is identical with (i), we conclude by virtue of Continuity Theorem for m.g.f. that (1) is the density of S_n.
Now let $Y = S_n/n = X/n$. $(X = S_n)$, then

$$F_Y(y) = P(Y \le y) = P(X \le ny) = \int_0^{ny} f_X(x)\, dx.$$

Differentiating : $f_Y(y) = nf_X(ny)$ which is the result (2).

Generalization. Let $Z_1, Z_2, \ldots, Z_n$ be i.i.d. Unif (a, b) variates, $Z = Z_1 + \ldots + Z_n$, $T = Z/n = \overline{Z}_n$ and let $h = b - a$. The p.d.f. of Z and that of T is

$$f_Z(z) = \frac{1}{h\Gamma(n)} \sum_{r=0}^{n} (-1)^r \binom{n}{r} [C_{(z-na)/h-r}]^{n-1}, \qquad na + rh < z \le na + (r+1)h. \qquad \ldots(3)$$

$$f_T(t) = \frac{n}{h\Gamma(n)} \sum_{r=0}^{n} (-1)^r \binom{n}{r} [C_{n(t-a)/h-r}]^{n-1}, \qquad a + \frac{rh}{n} < t \le a + \frac{(r+1)\,h}{n}. \qquad \ldots(4)$$

Proof. Let $X_i = (Z_i - a)/h$; then $X_i \sim$ Unif (0, 1). Now

$$X = X_1 + \ldots + X_n = [(Z_1 + \ldots Z_n) - na]/h = (Z - na)/h$$

$$F_Z(z) = P\{Z \le z\} = P\{hX + na \le z\} = P\{X \le (z-na)/h\} = \int_0^\theta g_n(x)\, dx. \; \theta = (z-na)/h$$

Differentiating w.r.t. z we obtain

$$f_z(z) = \frac{1}{h} g_n\left(\frac{z-na}{h}\right) = \frac{1}{h} \sum_{r=0}^{n} (-1)^r \binom{n}{r} [C_{(z-na)/h-r}]^{n-1} \frac{1}{\Gamma(n)}, \quad r < \frac{z-na}{h} \le r+1.$$

$$F_T(t) = P\{T \le t\} = P\{Z \le nt\} = \int_0^{nt} f_z(z)\, dz.$$

Differentiating : $F_T'(t) = n f_z(nt)$, which is the result (4).

18-41. *Special Cases Illustrated :*

1. $$h_2(y) = \frac{2}{\Gamma(2)} \sum_{r=0}^{2} (-1)^r \binom{2}{r} [C_{2y-r}]^2 \qquad \frac{r}{2} < y \le \frac{r+1}{2}$$ [Mean of two variates]

$$= 2[C_{2y-0} - 2C_{2y-1} + C_{2y-2}], [r = 0 \text{ gives } t_1 ; r = 1 \text{ gives } t_1 + t_2,\ r = 2 \text{ gives } t_1 + t_2 + t_3]$$

$$= 2\{2y\, I(0 < y \le 1/2) + [2y - 2(2y-1)]\, I(1/2 < y \le 3/2) + [2y - 2(2y-1) + 2y - 2]\, I(1 < y \le 3/2)\}$$

$$= \begin{cases} 4y, 0 < y \le 1/2 \\ 4(1-y),\ 1/2 < y \le 3/2 \\ 0,\ y > 3/2. \end{cases}$$

2. $$h_3(y) = \frac{3}{2} \sum_{r=0}^{3} (-1)^r \binom{3}{r} [C_{3y-r}]^2. \qquad \frac{r}{3} < y \le \frac{r+1}{3}$$ [Mean of three variates]

$$= \frac{3}{2} \begin{cases} (3y)^2, \quad 0 < y \le 1/3 & r = 0 \\ (3y)^2 - 3(3y-1)^2,\ \ 1/3 < y \le 2/3 & r = 1 \\ (3y)^2 - 3(3y-1)^2 + 3(3y-2)^2,\ \ 2/3 < y \le 1 & r = 2 \\ (3y)^2 - 3(3y-1)^2 + 3(3y-2)^2 - (3y-3)^2,\ \ y > 1 & r = 3 \end{cases}$$

$$= (27/2)\, y^2\, I(0 < y \le 1/3) + 27\, [(1/12) - (y - 1/2)^2]\, I(1/3) < y \le 2/3) + (27/2)\,(1-y)^2\, I(2/3 < y \le 1)$$

3. $$g_3(x) = \frac{1}{2} \sum_{r=0}^{3} (-1)^r \binom{3}{r} [C_{x-r}]^2, \quad r < x \le r+1.$$ [Sum of three variates]

$$= \frac{1}{2} \begin{cases} x^2, \quad 0 < x \le 1 & r = 0 \\ x^2 - 3(x-1)^2,\ 1 < x \le 2, & r = 1 \\ x^2 - 3(x-1)^2 + 3(x-2)^2,\ \ 2 < x \le 3, & r = 2 \\ x^2 - 3(x-1)^2 + 3(x-2)^2 - (x-3)^2, & x > 3 \end{cases}$$

$$= (1/2)x^2\, I(0 < x \le 1) + (3x - x^2 - 3/2)\, I(1 < x \le 2) + [(x-3)^2/2]\, I(2 < x \le 3) + 0\, I(x > 3).$$

Note. Some authors write $C(x)$ instead of C_x and some include left ends instead of right ends. For instance :

$$C(x) - 2\,C(x-1) + C(x-2) = 0\, I(x < 0) + x\, I\,(0 \le x < 1) + [x - 2(x-1)]\, I(1 \le x < 2)$$
$$+ [x - 2(x-1) + (x-2)]\, I(2 \le x < 3)$$

18-42. Bevy of Worked-out Problems

***Example* 1.** Let X, Y, Z be i.i.d. Unif(0, 1/2) variates. If $V = \max\{X, Y, Z\}$ and $W = \min\{X, Y, Z\}$, find $P\{1/3 < W$ and $V \le 1/2\}$; $P\{V \le 1/2\}$ and $P\{W > 1/3\}$.

Solution. Here $F(a) = \int_0^a dx = a_1$, so $F(1/2) = 1/2$, $F(1/3) = 1/3$. Now

$$\{V \le \tfrac{1}{2}\} = \{X \le \tfrac{1}{2}, Y \le \tfrac{1}{2}, Z \le \tfrac{1}{2}\} \Rightarrow P(V \le \tfrac{1}{2}) = [P(X \le \tfrac{1}{2})]^3 = [F(1/2)]^3 = 1/8$$

$$\{W > 1/3\} = \{X > 1/3, Y > 1/3, Z > 1/3\} \Rightarrow P(W > 1/3) = \{P(X > \tfrac{1}{3})\}^3 = [1 - F(1/3)]^3 = 8/27$$

$$\{1/3 < W, V \le 1/3\} = \{1/3 < X \le 1/2,\ 1/2 < Y \le 1/2,\ 1/2 < Z \le 1/2\}$$

$$\therefore\ P\{1/3 < W, V \le 1/3\} = [P(1/3 < X \le 1/2)]^3 = [F(1/2) - F(1/3)]^3 = (1/2 - 1/3)^3 = 1/216.$$

***Example* 2.** Let X and Y be i.i.d. Unif (a, b) variates and $k \in (a, b)$.

(i) Find the number k such that the probability that at least one of X and Y exceeds k is p.

(ii) Find the number k such that the probability that both X and Y are less than k equals the probability that exactly one of two is less than k.

Solution. Let $A = \{X > k\}$, $B = \{Y > k\}$; then

$$P(A) = P(B) = \int_k^b \frac{dx}{b-a} = \frac{b-k}{b-a};\ P(\bar{A}) = \frac{k-a}{b-a}. \qquad \text{...(1)}$$

(i) $p = P(A \cup B) = 1 - P(\bar{A})\,P(\bar{B}) = 1 - [(k-a)/(b-a)]^2$.

$$\therefore \qquad (k-a)/(b-a) = \sqrt{1-p} \Rightarrow k = a + (b-a)\sqrt{1-p}.$$

(ii) $$P(\overline{AB}) = P(\bar{A}\,\Delta\,\bar{B}) \Rightarrow P(\bar{A})\,P(\bar{B}) = P(\bar{A}) + P(\bar{B}) - 2P(\bar{A})\,P(\bar{B}).$$

Thus $$3P(\bar{A})\,P(\bar{B}) = 2P(\bar{A}) \Rightarrow P(\bar{B}) = (2/3),\ P(\bar{A}) \ne 0.$$

Using (1), this gives $k - a = (2/3)\,(b - a)$ or $k = a + [2(b - a)/3]$.

***Example* 3.** If a, b, c are randomly chosen between 0 and 1, find the probability that the quadratic equation $ax^2 + bx + c = 0$ has real roots.

Solution. The variates a, b, c are i.i.d. $U(0, 1)$ and their joint p.d.f. is

$$f(a, b, c) = 1,\ 0 \le a, b, c \le 1. \qquad [\because\ f(a, b, c) = f_1(a)\,f_2(b)\,f_3(c) = 1]$$

The roots of equation $ax^2 + bx + c = 0$ are real iff $b^2 \ge 4ac$. Thus,

$$P\{\text{roots real}\} = P\{b^2 \ge 4ac\} = 1 - P\{b^2 \le 4ac\} \qquad \text{...(1)}$$

$$P\{b^2 \le 4ac\} = \int_{a=0}^1 \int_{c=0}^1 \int_{b=0}^{\sqrt{4ac}} da\,db\,dc = \int_0^1 \int_0^1 \sqrt{4ac}\,da\,dc = 2\int_0^1 \sqrt{a}\,da \int_0^1 \sqrt{c}\,dc = \frac{8}{9}.$$

Thus, P {roots real} $= 1 - (8/9) = 1/9$. [by (1)]

***Example* 4.** If X is a random variable with a continuous distribution function F, then $Y = F(X)$ is $U(0, 1)$. Further, if $f(x) = \frac{1}{2}(x-1), 1 \le x \le 3$; $f(x) = 0$, otherwise determine what interval of Y corresponds to interval $1.1 \le X \le 2.9$.

Solution. $Y = F_X(x)$ is the c.d.f. X, then $Y \sim U(0, 1)$, [§ 8-16 (10)]

Further, for given p.d.f. : $F(x) = \int_1^x \frac{1}{2}(t-1)\,dt = \frac{1}{4}(x-1)^2, \quad 1 < x < 3$

Now $Y = F(X)$ is Unif (0, 1) ; so we define the mapping $y = \frac{1}{4}(x-1)^2$, then

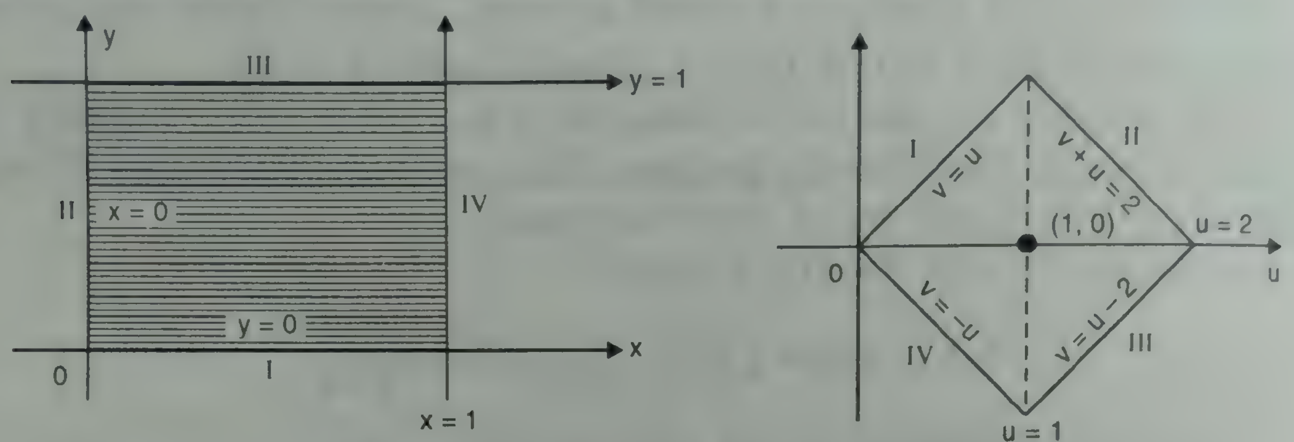
$$1.1 \le X \le 2.9 \Rightarrow \tfrac{1}{4}(1.1-1)^2 \le \tfrac{1}{4}(X-1)^2 \le \tfrac{1}{4}(2.9-1)^2, \Rightarrow 0.0025 < Y < 0.9025.$$

***Example* 5.** If X and Y are independent variates, each uniformly distributed over (0, 1), find the distribution of $U = X + Y$ and $V = X - Y$.

Solution. Let $u = x + y$, $v = x - y$, and invert them to get $x = \frac{1}{2}(u+v)$, $y = \frac{1}{2}(u-v)$.

$$\therefore \qquad |J| = |\partial(x, y)/\partial(u, v)| = \tfrac{1}{2}, \text{ so } dx\,dy = \tfrac{1}{2}du\,dv.$$

The range in the x-y plane and that in the u-v plane are as shown.



Mapping flat square to a Diamond square.

I : $y = 0, 0 \le x \le 1 \Rightarrow v = u, \quad 0 \le u + v \le 2,$ i.e. $0 \le u, v \le 1$.
II : $x = 0, 0 \le y \le 1 \Rightarrow v = -u, 0 \le u - v \le 2,$ i.e. $0 \le u, -v \le 1$.
III : $y = 1, 0 \le x \le 1 \Rightarrow u - v = 2, \quad 0 \le u + v \le 2,$ i.e. $1 \le u \le 2, -1 \le v \le 0$.
IV : $x = 1, 0 \le y \le 1 \Rightarrow u + v = 2, \quad 0 \le u - v \le 2,$ i.e. $1 \le u \le 2, 0 \le v \le 1$.

Since X and Y are indep.; their joint elemental p.d.f. is

$$dP_1(x, y) = f(x, y)\,dx\,dy = dx\,dy \Rightarrow dP_2(u, v) = \tfrac{1}{2}du\,dv. \qquad [f_1(x) = 1 = f_2(y)] \qquad \ldots(1)$$

We now integrate **out** v from Eq. (1) and obtain

$$\left.\begin{aligned} dP_3(u) &= \tfrac{1}{2}du\int_{-u}^{u} dv, \quad 0 \le u \le 1 \\ &\tfrac{1}{2}du\int_{u-2}^{2-u} dv, \quad 1 \le u \le 2 \end{aligned}\right] \Rightarrow g_1(u) = \begin{cases} u, & 0 \le u \le 1; \\ 2-u, & 1 \le u \le 2 \end{cases}$$

We now integrate out u from Eq. (1) and obtain

$$\left.\begin{aligned} dP_4(v) &= \tfrac{1}{2}dv\int_{v}^{2-v} du, \quad 0 \le v \le 1 \\ &\tfrac{1}{2}dv\int_{-v}^{v+2} du, \quad -1 \le v \le 0 \end{aligned}\right] \Rightarrow g_2(v) = \begin{cases} 1+v, & -1 \le v \le 0 \\ 1-v, & 0 \le v \le 1 \end{cases}$$

***Example* 6.** If $X_1, X_2, \ldots, X_n$ are i.i.d Unif (0, 1) variates, find the distribution of their geometric mean $G = (X_1 \,.\, X_2 \ldots X_n)^{1/n}$.

Solution. If $X \sim U(0, 1)$, then $f_X(x) = 1$, $0 < x < 1$, and

$$M(t : -\ln X) = \mathrm{E}(e^{-t\ln X}) = \mathrm{E}(e^{\ln X^{-t}}) = \mathrm{E}(X^{-t}) = \int_0^1 x^{-t}\,dx = (1-t)^{-1}, \; [t \ne 1].$$

$\therefore\ M(t:-\ell n\, X_1-\ldots-\ell n\, X_n)=[M(t:-\ell n\, X_1)]^n=(1-t)^{-n}$. $\quad$ [X_j are i.i.d.]

i.e. $\quad M(t:V)=(1-t)^{-n},\ V=-\Sigma\,\ell n\, X_j=-\ell n(X_1\,.\,X_2\ldots X_n)=-\ell n\,(G)^n$.

This being the m.g.f. of gam (1, n), it follows that the p.d.f. of V is

$$f(v)\ =\ e^{-v}\,v^{n-1}/\Gamma(n),\ 0<v<\infty. \qquad \ldots(1)$$

$$F_G(w)\ =\ P\{G\le w\}=P\{e^{-V/n}\le w\}=P\{-(V/n)\le \ell n\, w\}=P\{V\ge -n\,\ell n\, w\}$$

$$=\ 1-P\{V\le -n\,\ell n\, w\}=1-\int_0^{-w_1} f(v)\,dv\quad [w_1=n\,\ell n\, w]$$

Using DUIS by w, we recover

$$f_G(w)\ =\ -f(-w_1)(-n/w)=(n/w)\,f(-n\,\ell n\, w)$$

$$=\ \left(\frac{n}{w}\right)\frac{e^{n\ell n w}(-n\,\ell n\, w)^{n-1}}{\Gamma(n)}=\frac{nw^{n-1}}{\Gamma(n)}(-n\,\ell n\, w)^{n-1},\ 0<w<1$$

It is trivially verified : $\int_0^1 f_G(w)\,dw=1$.

***Example* 7.** Let k be a random number from (0, 1) and let

$$T=\ 1+[\ell n\, k/\ell n\, q],\ 0<p<1,\ q=1-p$$

where bracket function $[a]$ denotes greatest integer $\le a$. Then $T\sim$ gem (p).

Solution. For all $n\ge 1$,

$$P\{T=n\}\ =\ P\left\{\left[\frac{\ell n\, k}{\ell n\, q}\right]=n-1\right\}=P\left\{n-1\le\frac{\ell n\, k}{\ell n\, q}<n\right\}$$

$$=\ P\{(n-1)\,\ell n\, q\ge \ell n\, k>n\,\ell n\, q\},\qquad [\ell n\, q<0,\text{ so inequality reverses}]$$

$$\equiv\ P\{\ell n\,(q)^{n-1}\ge \ell n\, k>\ell n\,(q)^n\},\qquad [\ell n\text{ is monotone}\uparrow\text{ function}]$$

$$=\ P\{q^{n-1}\ge k>q^n\}=P\{q^n<k\le q^{n-1}\}$$

$$=\ q^{n-1}-q^n\quad [k\text{ is unif }(0,1)]$$

$$=\ q^{n-1}(1-q)=pq^{n-1},\ n\ge 1.$$

This proves that $T\sim$ gem (p).

Comments. (i) To simulate a r.v. $T\sim$ gem (p), all that is needed is to choose a point k at random from (0, 1) and let $T=1+[\ell n\, k/\ell n\, q]$.

(ii) To simulate (generate) $X\sim$ NB*(n, p), recall $X=X_1+X_2+\ldots+X_n$ where X_j are i.i.d. gem (p). Hence, to simulate X, choose n independent random numbers $k_1, k_2, \ldots, k_n$ from (0, 1) $k_i\sim$ Unif (0, 1) and let

$$X\ =\ \sum_{i=1}^{n}\left\{1+\left[\frac{\ell n\, k_i}{\ell n\, q}\right]\right\}.$$

***Example* 8.** Let $f(x)=12x/(1+x)^5$, $0<x<\infty$, $g(x)=2/(1+x)^3$, $0<x<\infty$.

Show that $h(x)\le c$, when $c\ge 3/2$, where $h(x)=f(x)/g(x)$. Apply rejection method to simulate (i.e. generate) X from $f(x)$.

Solution. We convert infinite interval] 0, ∞ [to a finite interval] 0, 1 [by using $y=1/(1+x)$ so that $f(x)\,dx=\phi(y)\,dy$. This gives equivalent :

$$f(y) = 12y^2(1-y), 0 < y < 1,\ g(y) = 2y, 0 < y < 1,\ h(y) = 6y(1-y).$$

$$\therefore \quad h'(y) = 6(1-2y), h''(y) = -12, h'(y) = 0 \Rightarrow y = \tfrac{1}{2},\ h(\tfrac{1}{2}) = \tfrac{3}{2}.\ \text{Thus}$$

$$\max h(y) = 3/2 \text{ and it occurs at } y = 1/2. \text{ So } c = h(1/2) = 3/2.$$

$$h(y)/c = 4y(1-y).$$

The rejection procedure is as under :

***Step* 1.** Generate random numbers u_1 and u_2 from interval $]\,0, 1\,[$.

***Step* 2.** If $u_2 \le 4u_1(1-u_1)$, stop and set $Y = u_1$. Otherwise return to step 1.

The average number of times that step 1 will be performed is $c = 3/2$.

Problems with Solutions Provided at the End of the Text

1*. A man and a woman agree to meet at a certain place between 10 and 11 o'clock They agree that the one arriving first will wait t hours, $0 \le t \le 1$, for the other to arrive. Assuming that the arrival times are independent and uniformly distributed, find the probability that they meet.

2*. If X is Unif $(-b, b)$, determine b if $P(\,|X| > 2) = 3/4$.

3*. If X is Unif $(1, 2)$, find z such that $P(X > z + \mu_X) = 1/4$.

4*. Let X be Unif $(-1, 3)$. Compare the upper bound on $P\{\,|X - \mu| \ge 2\sigma\}$, obtained from Chebyshev's inequality with exact probability.

5*. Let $X_1, X_2, \ldots, X_n$ be i.i.d. Unif $(-c, c)$ variates and define $L = \min(X_1, X_2, \ldots, X_n)$, $M = \max(X_1, X_2, \ldots, X_n)$. Evaluate $P(L \le -b \text{ or } M \ge a)$ when $0 < a, b < c$.

6*. The variates a and b are independently and uniformly distributed in the intervals $[0, 6]$ and $[0, 9]$ respectively. Find the probability that $x^2 - ax + b = 0$ has two real roots.

7*. If X_1, X_2, X_3 are independent r.v.'s, each uniformly distributed over $(-\tfrac{1}{2}, \tfrac{1}{2})$, show that $E(X^4) = 13/80$, where $X = X_1 + X_2 + X_3$.

8*. Let $f_X(x) = 6x(1-x)$, $0 \le x \le 1$. (i) Find y as a function of x such that Y has p.d.f.

$$g(y) = 3(1-\sqrt{y}), 0 \le y \le 1.$$

(ii) Show that $Z = 3X^2 - 2X^3$ follows Unif (0, 1) law.

9*. Derive a random sample from the distribution with p.d.f. :

(i) $f(x) = 2xe^{-x^2},\ x > 0$; (ii) $f(x) = ra^r/x^{r+1},\ x \ge a$.

10*. Random variables X and Y are i.i.d. and have Unif (0, 1) distribution. Show that

$$\phi(t : |X - Y|) = 2[1 + it - e^{it}]/t^2, \qquad (t \text{ real}).$$

Deduce that $E[|X-Y|]^n = 2/(n+1)(n+2)$.

11*. If X and Y are i.i.d. Unif (0, 1) variates, find the distributions of XY and X/Y.

12*. If X, Y, Z are i.i.d. Unif (0, 1)-variates, show that the distribution of $W = 2(X + Y + Z) - 3$ is given by

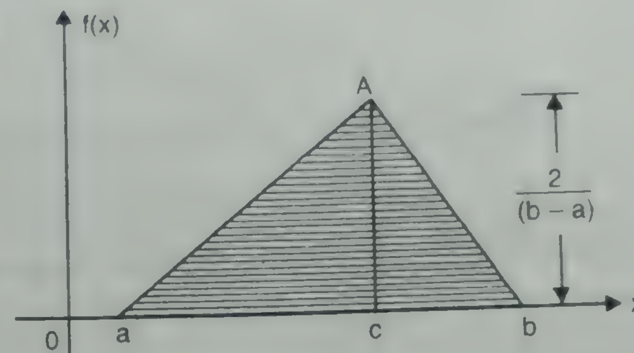
$$f(w) = \frac{(w+3)^2}{16}, -3 < w < -1;\ f(w) = \frac{(3-w^2)}{8}, -1 \le w \le 1,\ f(w) = \frac{(w-3)^2}{16}, 1 \le w \le 3 \quad \ldots(A)$$

Compare the p.d.f. of W with that of $N(0, 1)$ and show that $2W - 3$ is not a bad approximation for $N(0, 1)$.

18-50. *Triangular Distribution and its m.g.f.*

A random variable X possessing the following p.d.f. is called **triangular variate** :

$$f(x) = \begin{cases} 2(x-a)/(b-a)(c-a), & a < x < c \\ 2(b-x)/(b-a)(b-c), & c < x < b \\ 0, & \text{otherwise} \end{cases} \quad \ldots(1)$$



Such a variate is written Trg (a, b) ; there is peak at $x = c$.

$$M(t : X) = \int_a^c e^{tx} \frac{2(x-a)\,dx}{(b-a)(c-a)} + \int_c^b e^{tx} \frac{2(b-x)\,dx}{(b-a)(b-c)}$$

$$= \frac{2}{(b-a)(c-a)} \left\{ \frac{e^{at}}{t^2} + e^{ct}\left(\frac{c-a}{t} - \frac{1}{t^2} \right) \right\} + \frac{2}{(b-a)(b-c)} \left\{ \frac{e^{bt}}{t^2} - e^{ct}\left(\frac{b-c}{t} + \frac{1}{t^2} \right) \right\}$$

$$= \frac{2}{t^2} \left\{ \frac{e^{at}}{(a-b)(a-c)} + \frac{e^{bt}}{(b-a)(b-c)} + \frac{e^{ct}}{(c-a)(c-b)} \right\}, a < c < b. \qquad \ldots(2)$$

There is an excellent symmetry in this expression ; and the form seems to be new. If we expand the exponents involved, Coeff. of t^{-2} and t^{-1} disappear, constant is unity confirming $M(0) = 1$; and

$$\text{E}(X) = \frac{1}{3} \left\{ \frac{a^3}{(a-c)(a-b)} + \frac{b^3}{(b-a)(b-c)} + \frac{c^3}{(c-a)(c-b)} \right\}$$

$$\text{E}(X^2) = \frac{1}{6} \left\{ \frac{a^4}{(a-c)(a-b)} + \frac{b^4}{(b-a)(b-c)} + \frac{c^4}{(c-a)(c-b)} \right\}$$

these being the Coefficients of t and $t^2/2\,!$ in expression (2).

***Example* 1.** Find the m.g.f. for the triangular distribution

$$f(x) = x, \;\; 0 < x \le 1; \;\; f(x) = 2 - x, \;\; 1 < x \le 2; \;\; f(x) = 0, \text{otherwise}.$$

If Y and Z are independent $U(0, 1)$ variates, show that $Y + Z$ has triangular distribution.

Solution.

$$M(t) = \int_0^1 e^{tx}\, x\, dx + \int_1^2 (2-x)\, e^{tx}\, dx = \left[\frac{e^{tx}(xt-1)}{t^2} \right]_0^1 + \left[\frac{e^{tx}(1+2t-xt)}{t^2} \right]_1^2 = \frac{(e^t-1)^2}{t^2} \qquad \ldots(i)$$

Since Y and Z are independent $U(0, 1)$ variates, so

$$M(t : Y + Z) = M(t : Y) \,.\, M(t : Z) = [(e^t - 1)/t]^2. \qquad \ldots(ii)$$

From (i) and (ii) it follows that $V = Y + Z$ is Trgl $(0, 2)$ variate, with peak at $v = 1$.

$$g(v) = v, \; 0 \le v \le 1 \; ; \; g(v) = 2 - v, \; 1 \le v \le 2, \; g(v) = 0, \text{ otherwise.}$$

***Example* 2.** If X and Y are i.i.d. Unif $(-a, a)$ variates, find the density of $T = X + Y$.

Solution. Recall that the m.g.f. of $X \sim$ Unif $(-a, a)$ is $M(t : X) = (e^{at} - e^{-at})/2at$.

$$\therefore \quad M(t : X + Y) = M(t:X).M(t:Y) = (e^{at} - e^{-at})^2 / 4a^2t^2$$

$$= \frac{2}{t^2}\left\{\frac{e^{2at}}{(2a+2a)(2a-0)} + \frac{e^{0t}}{(0+2a)(0-2a)} + \frac{e^{-2at}}{(-2a-0)(-2a-2a)}\right\}.$$

This shows that $X + Y \sim$ Trgl $(-2a, 2a)$; hence the density of $T = X + Y$ is

$$g(u) = (2a+u)/4a^2, \ -2a < u \le 0; \quad g(u) = (2a-u)/4a^2, \ \ 0 < u \le 2a.$$

Note. The density of $X - Y$ also follows as under. Thus, if X and Y are i.i.d. $U(a, b)$. Then

$$M(t : X - Y) = M(t:X)\, M(t:-Y) = (e^{bt} - e^{at}) \times (e^{-bt} - e^{at}) / [-(b-a)^2 t^2]$$

For brevity, put $b - a = h$, then

$$M(t : X - Y) = (e^{ht} - 2 + e^{-ht})/h^2t^2 = \frac{2}{t^2}\left\{\frac{e^{ht}}{(h-0)(h+h)} + \frac{e^{0t}}{(0-h)(0+h)} + \frac{e^{-ht}}{(-h)(-2h)}\right\}$$

This being m.g.f. of Trgl $(-h, h)$; it follows that $V = X - Y$ has the density

$$g(v) = (v + h)/h^2, \quad -h < v < 0 \ ; \quad g(v) = (h - v)/h^2, \quad 0 < v < h.$$

Remark. If the densities of $X + Y$ and $X - Y$ are both needed, then Jacobian method may be tried.

Exercise 18(a)

1. A r.v. X uniformly distributed over an interval of unit length is such that $P(1/4 < X < 1/2) = 1/4$. Determine the interval. **[Ans.** Undetermined]
2. If X has the distribution : $f(x) = 5/A$, $-A/10 \le x \le A/10$; $f(x) = 0$, otherwise determine the constant A when $P\{\,|X| < 2\} = 2P\{\,|X| > 2\}$. **[Ans.** 30]
3. (*a*) For the rectangular distribution $dF = k\,dx$, $1 \le x \le 2$.
   Show : Arithmetic mean > Geometric Mean > Harmonic mean.

   (b) Let X .. $U(0, 1)$. Find Corr (X, Y) when $Y = X_n$. **[Ans.** $\rho = \sqrt{(6n+3)}/(n+2)$]
4. Suppose a string, 1 meter long, is to be cut in two at a random point, along its length. Let X be the point where cut occurs and let its p.d.f. be $f(x) = 1$, $0 < x < 1$.
   What is the probability that the longer piece is atleast twice the length of the shorter ?
5. If X be $U(-a, a)$, $a > 0$; determine a so that
   (a) $P(X > 1) = 0.3$, (b) $P(X \le 1/2) = 0.3$, (c) $P(\,|X| > 1) = P(\,|X| < 1)$. **[Ans.** 5/2, –5/2, 2]
6. If X is $U(-3, 3)$, find k such that $P(X < k) = 1/3$. Also compute $P(X = 2)$, $P(X < 2)$ and $P\{\,|X - 2| < 2\}$. **[Ans.** 0, 5/6, 1/2]
7. If X be $U(a, 9)$ and $P\{3 < X < 5\} = 2/7$, find a and compute $P\{\,|X - 5| < 2\}$. Find also a real number r such that $P\{X > r\} = 1/4$. **[Ans.** $a = 2$, $r = 29/4$]
8. If $M(t : X) = (e^{5t} - e^{4t})/t\,; t \ne 0$, $M_X(0) = 1$, find Var (X) and $P\{4.9 > X > 4.7\}$. **[Ans.** 1/2, 0.2]
9. If X is uniformly distributed with mean 1 and variance 4/3, find $P\{X < 0\}$. **[Ans.** 1/4]

10. If X and Y are i.i.d. $U(0, 1)$ variates, find

(a) $P(X^2 < 0.5)$, (b) $P(e^{-X} \le 0.5)$, (c) $P\{\cos(\pi X) \le 0.5\}$,

(d) $P\{X + Y \le 0.5\}$, (e) $P\{X - Y \le 0.5\}$, (f) $P\{XY \le 0.5\}$,

(g) $P\{(X/Y) \le 0.5\}$, (h) $P\{X^2 + Y^2 < 0.5\}$.

[**Ans.** $1/\sqrt{2}$, $1 - \ln 2$, $2/3$; (d) $1/8$, $7/8$; $(1 + \ln 2)/2$, $1/4$, $\pi/8$]

11. On the axis $n + 1$ points are taken independently between the origin and $x = 1$, all the positions being equally likely. Show that probability that $(k + 1)$th of these points, counted from the origin, lies in the interval $(x - (1/2)\,dx, x + (1/2)\,dx)$ is $(n + 1)\,{}^nC_k\, x^k (1 - x)^{n-k}\, dx$. Verify that the integral of this expression from $x = 0$ to $x = 1$ is unity.

12. A point-size worm is inside a spherical apple $x^2 + y^2 + z^2 = 4a^2$; its position is uniformly distributed. If the apple is eaten down to a core, determined by the intersection of the sphere and the cylinder $x^2 + y^2 = a^2$, find the chance that the worm will be eaten up. [**Ans.** $\sqrt{27}/8$]

13. The quadratic equation $x^2 - ax + b = 0$ has real roots X_1, X_2, $(X_1 > X_2)$ and $b > 0$ is an unknown quantity that can have a uniform distribution in the permissible range of variation. Find Var (X_1) and Var (X_2). [**Ans.** $\sigma_1^2 = \sigma_2^2 = a^2/72$]

14. (a) Both roots of equation $x^2 + ax + b = 0$ can take all values from -1 to $+1$ with equal probabilities. Obtain p.d.f. for the coefficients a and b.

[**Ans.** $f(a) = (2 - |a|)/4$, $g(b) = -(1/2)\ln|b|$]

(b) Let a and b be jointly, uniformly distributed over the quarter circle $a^2 + 1b^2 < 1$, $a > 0$, $b > 0$. Show that the probability that the equation $x^2 + 2\sqrt{a}\,x - (b - 1) = 0$ will have no real roots is $2/\pi$.

15. (a) Find the probability that the roots of equation $x^2 + 2bx + c = 0$ should be real, given that $b \sim U(-\alpha, \alpha)$ and $c \sim U(-\beta, \beta)$ are independent.

(b) In the equation $x^2 + 2ax + b = 0$; a and b independently are equally likely to assume values in the interval $[-1, 1]$. Find the probability that the roots of the equation are real.

16. Suppose that a, b, c are positive, independent variates with a common c.d.f. "F". Show that the quadratic $x^2 + bx + c = 0$ has real roots with probability $\int_0^\infty \int_0^\infty F(x^2/4y)\, dF(x)\, dF(x)$.

17. If X is $U(0, 1)$, then Y is $N(0, 1)$ where $X = (1/2)[1 + (2/\sqrt{2\pi}) \int_0^Y \exp(-t^2/2)\, dt]$. [**Ans.** 1/3]

18. (a) If $X \sim U(1, 4)$ find $P(Y < 0)$ where $Y = X^2 - 4$.

(b) If $X \sim U(0, 1)$, compare $F\{|X - \mu| < k\sigma\}$ with the values given by Chebyshev's inequality for $k = 5/4, 3/2, 7/2$ and 2.

19. (a) If $X \sim U(0, 1)$, find the distribution of $1/X$. Find $E(1/X)$ if it exists.

(b) If X is $U(-\pi/2, \pi/2)$, find the p.d.f. of $Y = \tan X$.

[**Ans.** (a) $f(y) = y^{-2}$, $1 \le y < \infty$, $E(Y) \to \infty$, (b) $1/\pi\,(1 + y^2)$]

20. (i) If $X \sim U(a, b)$, show that $cX + d \sim U(ca + d, cb + d)$.

(ii) If X is $U(0, 1)$, show that $Y = (b - a)X + a$ is $U(a, b)$.

21. If $X \sim U(-1, 1)$, find p.d.f. of X^2. Further, if $X \sim U(0, 1)$, find the p.d.f. of $\sqrt{X}, X^4, X^{-1}, e^X, -\ln X, X/(1 + X), -\lambda^{-1}\ln(1 - X), (\lambda > 0)$.

22. (a) Let $X \sim U(0, 1)$ and $Y \sim U(0, X)$. Find the joint density of X and Y and the marginal density of Y.

(b) If X is $U(0, 1)$ and Y is related to X by $X = \int_{-\infty}^{Y} f(t)\,dt$, $\int_{-\infty}^{\infty} f(t) = 1, f(t) \geq 0$, show that $f(t)$ is the density of Y.

23. Two independent variates are each uniformly distributed within the range $-a$ to $+a$. Show that their sum X has a p.d.f. given by

$$f(x) = (2a + x)/4a^2, \quad -2a \leq x \leq 0\,; \quad f(x) = (2a - x)/4a^2, \quad 0 < x \leq 2a.$$

Verify that the m.g.f. calculated from the value of $f(x)$ is equal to $[(\sinh at)/at]^2$.

24. Let $X_1, X_2, ..., X_n$, be i.i.d. $U(0, 1)$ random variables. If $Y_1 \sim U(0, 1)\,; Y_2 \sim U(0, Y_1), ..., Y_n \sim U(0, Y_{n-1})$, show that $Y_1 = X_1, Y_2 = X_1 . X_2 ... Y_n = X_1, X_2, ..., X_n$.

If Z is the number of $Y_1, Y_2, .., Y_n$ in $[t, 1], 0 < t < 1$, show that Z is Pois $(-\ln t)$.

25. If X, Y, Z, are i.i.d. $U(0, 1)$ variates, obtain the distributions of $\frac{1}{2}(X + Y)$, and $\frac{1}{3}(X + Y + Z)$.

[**Ans.** §18-41]

26. Let X and Y be i.i.d. $U(0, 1)$ variates. Define

$Z = X + Y$, if $0 \leq X + Y \leq 1\,; Z = X + Y - 1$, if $1 \leq X + Y \leq 2$. Show that Z is $U(0, 1)$.

27. Let X be a variate for which $E(X) = \mu$ and Var $(X) = \sigma^2$. Further, let Y be $U(a, b)$ variate. Determine a and b such that $E(X) = E(Y)$ and Var $(X) =$ Var (Y).

28. Let X and Y be i.i.d. $U(0, 1)$ variates. Find the conditional expectation of $(X + Y)^2$ given $(X - Y)$.

29. Let $X \sim U(0, 10)$ and define Y by $Y = X^2$, if $0 \leq X \leq 6$, $Y = 3$, if $6 \leq X \leq 10$.

Find the conditional expectation of Y given $2 \leq Y \leq 4$.

30. Let $X_1, X_2, ..., X_n$ be a random sample from $U(0, a)$ distribution. Find the c.d.f. of $T = \min \{X_1, X_2, ..., X_n\}$. [**Ans.** $f(t) = (n/a)[1 - (t/a)]^{n-1}, 0 < t < a$]

31. If X and Y are i.i.d. $U(0, 1)$ variates, find the p.d.f. of (i) X/Y^2, (ii) max (X, Y)/min (X, Y).

32. Let X and Y be i.i.d. $U(0, 1)$ variates. Let $U = XY$ and $V = (1 - X)(1 - Y)$. Find (i) the joint p.d.f. of U and V, (ii) the marginal distributions of U and V, (iii) the conditional distribution of U, given $V = v$.

33. Let $X_1, X_2, ..., X_n$ be i.i.d. $U(0, 1)$ variates. Show that the p.d.f. of $X = X_1, X_2, ..., X_n$ is $f(x) = (-\ln x)^{n-1}/(n - 1)\,!, \ 0 < x \leq 1\,; f(x) = 0$, otherwise.

34. If $X_1, X_2, ..., X_n$ and Z are i.i.d. $U(0, 1)$ variates, find the p.d.f. of $X_1, X_2, ..., X_n/Z$.

35. Let $X_1, ..., X_n$ be mutually indep. variates ; $X_k \sim U(k, k + 1), k = 1, 2, .., n$. Find the p.d.f. of $X_1 + X_2 + ... + X_n - n(n + 1)/2$.

36. Let $X \sim U(0, 1)$. For the variate $Z_\lambda = \lambda^{-1} \{X^\lambda - (1 - X)^\lambda]$, find β_1 and β_2 [Z_λ is Tukey's symmetrical lambda variate].

37. Let X be a variate ranging over $[0, 1]$. If $P\{x < X \leq y\}$ depends on the length $y - x$, for all $0 \leq x < y \leq 1$, then X is $U(0, 1)$.

38. Let X_k, $1 \leq k \leq n$ be indep. variates such that X_k is $U(-a_k, a_k)$. If $|a_k| < a$ and $\Sigma a_k^2 \to \infty$; show that Linderberg-Feller condition is satisfied and hence the C.L.T. holds.

39. Let X and Y be independent variates such that $E(|X|) < \infty$ and Y is $U(0, a)$. If $Z = a\,\{(X + Y)/a]$, then $E(X) = E(Z)$. ($[k]$ denotes the greatest integer contained in k).

40. Let X be a non-negative variate of continuous type. Let $[X]$ be Pois (λ) and $X - [X]$ be $U(0, 1)$. If $[X]$ and $X - [X]$ are assumed independent, find the p.d.f. and variance of X. ($[x]$ is integer contained in x and $x - [x]$ is fractional part of x).

41. Suppose (i) $f(x) = cg(x) \cdot H(x)$, (ii) $f(x) = cg(x)[1 - H(x)]$ where f, g are p.d.f's. and H is the c.d.f. of a variate with p.d.f. : $h(\cdot)$. A variate Y is chosen from density $h(\,.\,)$ and an independent value X is chosen from density $g(\cdot)$. Show that if x is accepted iff (i) $y < x$ [case (ii) $y > x$], then the accepted X has the distribution $f(\cdot)$.

EXPONENTIAL DISTRIBUTION

18-60. *Definition*

The continuous variate which is distributed according to the probability law

$$f(x) = \lambda e^{-\lambda x}, (\lambda > 0), 0 < x < \infty; f(x) = 0; \text{ otherwise} \qquad ...(1)$$

is called the *exponential* variate with parameter λ, f is called the exponential density (function). Any variate possessing the p.d.f. (1) is expressed as $X \sim$ Expo (λ).

Note. To treat mean $1/\lambda$ (= β say) as parameter, sometimes density is written

$$f(x) = (1/\beta) \exp(-x/\beta).$$

Comment. Some authors call (1) as *negative* exponential distribution. The term *negative* is really redundant since $\int_0^\infty e^x \, dx \to \infty$.

Cumulative Distribution Function :

$$F(x) = \int_0^x \lambda e^{-\lambda t} \, dt = 1 - e^{-\lambda x}, x > 0; \quad F(x) = 0, \text{otherwise.}$$

$$F(k) = P(X \le k) = 1 - e^{-\lambda k}; P(X \ge k) = e^{-\lambda k} \quad \text{[Survival function].}$$

18-61. *Simple Moments. M.G.F. Cumulant Generating Function*

$$\mu_r' = \mathrm{E}(X^r) = \int_0^\infty \lambda e^{-\lambda x} x^r \, dx = \frac{\Gamma(r+1)}{\lambda^r} = \frac{r!}{\lambda^r}.$$

Thus, $\quad \mu_1' = 1/\lambda, \mu_2' = 2/\lambda^2, \mu_2 = \mu_2' - \mu_1'^2 = 1/\lambda^2.$

If follows that for Expo (λ) : Mean <, = , > variance $\Leftrightarrow \lambda >=, < 1$.

$$M(t : X) = [1 - (t/\lambda)]^{-1}. \qquad [\S\ 8\text{-}16\ (13)]$$

Characteristics Function : $\phi(t : X) = [1 - (it/\lambda)]^{-1}$.

Expanding the Binomial expression involved in m.g.f. we get

$$M(t : X) = \sum_{r=0}^{\infty} \left(\frac{t}{\lambda}\right)^r = \sum_{r=0}^{\infty} \left(\frac{r!}{\lambda^r}\right) \cdot \frac{t^r}{r!}; \Rightarrow \mu_r' = r!/\lambda^r$$

$$K(t : X) = \ln M(t : X) = -\ln[1 - (t/\lambda)] = \sum_{r=1}^{\infty} \frac{(t/\lambda)^r}{r} = \sum_{r=1}^{\infty} \frac{(r-1)!}{\lambda^r} \frac{t^r}{r!}$$

$\therefore \quad k_r = (r-1)!/\lambda^r. \quad$ [Coeff. of $t^r/r!$]

Thus, $\quad k_1 = 1/\lambda, \; k_2 = 1/\lambda^2, \; k_3 = 2/\lambda^3, \; k_4 = 6/\lambda^4, \; ...$

Pearson Coefficient, Median and Mode

Since $\mu = k_1 = 1/\lambda, \; \mu_2 = k_2 = 1/\lambda^2, \mu_3 = k_3 = 2/\lambda^3, \; \mu_4 = k_4 + 3k_2^2 = 9/\lambda^4$.

$\therefore \quad \beta_1 = \mu_3^2/\mu_2^3 = 4; \; \beta_2 = \mu_4/\mu_2^2 = 9; \gamma_1 = \sqrt{\beta_1} = 2; \; \gamma_2 = \beta_2 - 3 = 6.$

$\gamma_1 > 0$ implies that Exponential distribution is positively skewed and $\gamma_2 > 0$ means the distribution is Leptocurtic. Note that the point ((β_1, β_2) lies on the st. line $\beta_1 - \beta_2 - 3 = 2$.

Median. If m is the median of exponential distribution then

$$\frac{1}{2} = \int_0^m \lambda e^{-\lambda x}\,dx = 1 - e^{-\lambda m} \Rightarrow e^{-\lambda m} = \frac{1}{2}, \quad \text{i.e.} \quad m = \frac{\ell n\, 2}{\lambda}.$$

Mode.
$$f(x) = \lambda e^{-\lambda x}, f'(x) = -\lambda^2 e^{-\lambda x},\ f''(x) = \lambda^3 e^{-\lambda x}$$

$f'(x) = 0 \Rightarrow e^{-\lambda x} = 0$, which is not possible. Hence expo(λ) possesses no modal value.

18-62. *Shifted Exponential Distribution ($X \sim \text{Expo}_\theta\ (\lambda)$)*

$$f(x) = \lambda e^{-\lambda(x-\theta)},\ x \geq \theta,\ \lambda > 0;\ f(x) = 0, \text{ elsewhere.}$$

$$M(t : X) = \mathsf{E}(e^{tX}) = \lambda \int_\theta^\infty e^{-\lambda(x-\theta)}.e^{tx}\,dx = \lambda e^{t\theta} \int_0^\infty e^{-(\lambda - t)y}\,dy, \qquad [y = x - \theta]$$

$$= \lambda e^{\theta t}/(\lambda - t) = e^{\theta t}\,[1 - (t/\lambda)]^{-1}.$$

18-63. *Memory-less Property or Markov Property*

A variate X is exponentially distributed iff it has no memory in the sense that

$$P\{X > x + h \mid X > h\} = P\{X > x\}, \quad h \geq 0, x > 0. \qquad ...(1)$$

Proof. Firstly we observe that if $X \sim$ Expo (λ), then **Reliability function** or **Survival function** is

$$R(t) \equiv P\{X > k\} = \int_k^\infty \lambda e^{-\lambda x}\,dx\ e^{-\lambda k}.$$

$$\therefore \quad P\{X > x + h \mid X > h\} = \frac{P\{X > x + h\}}{P\{X > h\}} = \frac{e^{-\lambda(x+h)}}{e^{-\lambda h}} = e^{-\lambda x} = P\{X > x\}.$$

Thus, X lacks memory in the sense of Eq. (1). A statement equivalent to (1) is

$$P\{X > x + h\} = P\{X > x\}\,P\{X > h\}. \qquad ...(2)$$

Converse. We assume that (2) is true, and then show that X is Exponential r.v. Let

$$R(x) = P\{X > x\} \qquad [R(0) = P(X > 0) = 1]$$

$$\therefore \quad R(x + h) = P\{X > x + h\} = R(x)\,R(h), \qquad [\text{by (2)}]$$

So
$$\frac{R(x+h) - R(x)}{h} = \frac{R(x)\,[R(h) - 1]}{h} = \frac{R(x)\,[R(h) - R(0)]}{h}$$

Let $h \to 0$ to conclude that

$$R'(x) = R(x)\,R'(0) \text{ or } (dR/R) = -\lambda dx \qquad [\lambda = -R'(0)]$$

The solution of this differential equation is $R(x) = Ae^{-\lambda x}$.

Since $A = R(0) = 1$, it follows that the distribution (i.e. anti-survival function) of X is

$$F_X(x) = 1 - R(x) = 1 - e^{-\lambda x},\ x > 0 \Rightarrow F_X' = f_X(x) = \lambda e^{-\lambda x}\ ; x > 0.$$

Observe that $R'(0) = -F'(0) < 0$, which provides $\lambda > 0$.

Remark. Expo (λ) distribution is a continuous analogue of *Non-ageing* discrete Geom (p).

18-70. Convolution of Exponential Variates

Let $X_j \sim$ expo (λ_j), $j = 1, 2$ be independent variates $(\lambda_1 \neq \lambda_2)$. The convolution (statistical sum) formula gives

$$\begin{aligned} f_{X_1+X_2}(t) &= \int_0^t f_1(z)\, f_2(t-z)\, dz = \int_0^t \lambda_1 e^{-\lambda_1 z}.\lambda_2 e^{-\lambda_2(t-z)} dz \\ &= \lambda_1\lambda_2 e^{-\lambda_2 t} \int_0^t e^{-(\lambda_1-\lambda_2)z}\, dz = \frac{\lambda_1\lambda_2 e^{-\lambda_2 t}}{(\lambda_1-\lambda_2)}[1-e^{-(\lambda_1-\lambda_2)t}] \\ &= \frac{\lambda_1}{\lambda_1-\lambda_2}(\lambda_2 e^{-\lambda_2 t}) + \frac{\lambda_2}{(\lambda_2-\lambda_1)}(\lambda_1 e^{-\lambda_1 t}) \\ &= \frac{\lambda_2}{\lambda_2-\lambda_1} f_1(t) + \frac{\lambda_1}{\lambda_1-\lambda_2} f_2(t) \qquad \text{...(1)} \end{aligned}$$

Similarly, if $X = X_1 + X_2 + X_3$, then

$$f_X(t) = f(t_1)\frac{\lambda_2\lambda_3}{(\lambda_2-\lambda_1)(\lambda_3-\lambda_1)} + f(t_2)\frac{\lambda_3\lambda_1}{(\lambda_1-\lambda_2)(\lambda_3-\lambda_2)} + f(t_3)\frac{\lambda_1\lambda_2}{(\lambda_2-\lambda_3)(\lambda_1-\lambda_3)}.$$

The general formula for $S = X_1 + X_2 + \ldots + X_n$ is

$$f_S(t) = \sum_{i=1}^{n} \frac{\lambda_1.\lambda_2\ldots\lambda_{i-1}\lambda_{i+1}\ldots\lambda_n}{(\lambda_1-\lambda_i)\ldots(\lambda_{i-1}-\lambda_i)(\lambda_{i+1}-\lambda_i)\ldots(\lambda_n-\lambda_i)}.f_i(t)$$

The variate $S = X_1 + X_2 + \ldots + X_n$ is called *hypo exponential variate.*

Note. $\quad I = \int_0^\infty f_S(t)\,dt = \sum_{i=1}^{n} \frac{\lambda_1\ldots\lambda_{i-1}\lambda_{i+1}\ldots\lambda_n}{(\lambda_1-\lambda_i)\ldots(\lambda_{i-1}-\lambda_i)\ldots(\lambda_n-\lambda_i)} = \sum_{i=1}^{n} C_{i,n}$ (say).

Here $C_{i,n}$ are not probabilities, since some of them could be negative.

18-80. Relation between Exponential and Uniform Distributions

(i) If $X \sim$ Expo (λ), then $Y = e^{-\lambda X}$ is Unif (0, 1).

(ii) If X is Unif(0, 1), then $Y = -(1/\lambda)\ln X$ is expo (λ).

Proof. Here $dP_1(x) = \lambda e^{-\lambda x}\,dx,\ 0 \le x < \infty$; put $y = e^{-\lambda x}$, $dy = -\lambda e^{-\lambda x}\,dx$, $0 \le x < \infty \Rightarrow 0 \le y < 1$; hence reversing sign

$$dP_2(y) = 1\,dy,\ 0 \le y \le 1;\ \text{ i.e. } \quad Y \sim U(0,1).$$

$$M_y(t) = \mathrm{E}(e^{tY}) = \mathrm{E}\{e^{-(t/\lambda)\ln X}\} = \mathrm{E}(X^{-t/\lambda}) = \int_0^1 x^{-t/\lambda}\,dx = \left(1-\frac{t}{\lambda}\right)^{-1} \Rightarrow Y \sim \text{expo}(\lambda).$$

Comments. If X is *shifted* exponential [$X \sim$ Expo (λ) : $f(x) = \lambda e^{-\lambda(x-\theta)}, x \ge \theta$], then $Y = e^{-\lambda(X-\theta)}$ is $U(\theta, 1)$.

18-81. Distribution of Minimum of Expo-variates

Let $X_i \sim$ expo (λ_i), $1 \le i \le n$ be indep. variates and $Z = \min\{X_1, \ldots, X_n\}$. Then $Z \sim$ Expo $(\sum\lambda_i)$.

Proof. We have $Z = \min\{X_1, \ldots, X_n\}$. To obtain the p.d.f. of Z, fix a real number $c > 0$. Then

$$P\{Z \geq c\} = P\{X_1 \geq c, X_2 \geq c, \ldots, X_n \geq c\} \qquad \text{[def. of min } (x_1, \ldots, x_n)]$$
$$= P\{X_1 \geq c\}. P\{X_2 \geq c) \ldots P(X_n \geq c\}$$
$$= e^{-\lambda_1 c}.e^{-\lambda_2 c} \ldots e^{-\lambda_n c} = e^{-(\lambda_1 + \ldots + \lambda_n)c}.$$

This is the survival (or Reliability) function of expo $(\Sigma\lambda_i)$. Thus

$$Z \sim \text{expo } (\Sigma\lambda_i) \text{ and } \mathsf{E}\{Z\} = 1/(\lambda_1 + \lambda_2 + \ldots + \lambda_n).$$

Cor. If $X_1, X_2, \ldots, X_n$ are i.i.d. Expo (β), and $Z = \min\{X_1, \ldots, X_n\}$ then $Y = nZ$ is Expo (β)

Proof. From (1), the elemental probability differential for Z, when $\lambda_i = \beta$, $1 \leq i \leq n$ is

$$dP_1(z) = n\beta e^{-n\beta z}\, dz, \; 0 < z < \infty.$$

Let $y = nz$, so that $dy = n\, dz$, and this result reduces to

$$dP_2(y) = \beta e^{-\beta y}\, dy, \; 0 \leq y < \infty, \; \Rightarrow \; Y \sim \text{Expo}(\beta).$$

18-82. *Bevy of Worked-out Problems*

Example 1. If $X \sim$ expo (α) and $Y \sim$ expo (β) are independent then $Z = \min\{X, Y\}$ is expo $(\alpha + \beta)$. Further, if variate W is defined by $W = 1$ when $X \leq Y$, and $W = 0$ when $X > Y$, then $\mathsf{E}(W) = \alpha/(\alpha + \beta)$ and the variates W and Z are mutually independent.

Solution. We already know that $Z \sim$ expo $(\alpha + \beta)$. Now by Multistage E-Rule

$$\mathsf{E}(W) = \mathsf{E}(1)\, P\{X \leq Y) + \mathsf{E}(0)\, P(X > Y) = P(X \leq Y).$$

To find $P(X \leq Y)$, use $P(B) = \Sigma(P(B \mid A_j)\, P(A_j)$ so that

$$P(X \leq Y) = \int_0^\infty P(X \leq Y \mid X = x)\, \alpha e^{-\alpha x}\, dx \quad [P(Y \geq x) = e^{-\beta x}]$$
$$= \int_0^\infty \alpha\, e^{-(\alpha+\beta)x}\, dx = \frac{\alpha}{\alpha + \beta}.$$

Thus $$\mathsf{E}(W) = \alpha/(\alpha + \beta).$$

Example 2. Let $X \sim$ expo (λ). Find $\mathsf{E}[X \mid (X \geq t)]$.

Solution. If $t < 0$, $\{X \geq t\} = \Omega$,

$$\therefore \qquad \mathsf{E}\{X \mid X \geq t\} = \mathsf{E}\{X \mid \Omega\} = \mathsf{E}(X) = 1/\lambda$$

as $\mathsf{E}(X)$ is reciprocal of parameter λ.

Now let $t > 0$. Then, let $B = \{X \geq t\}$

$$P(B) = P(X \geq t) = \int_t^\infty \lambda e^{-\lambda x}\, dx = e^{-\lambda t}.$$

$$\mathsf{E}(X \mid B) = \mathsf{E}(I_B X)/P(B). \qquad \text{[def.]} \qquad \ldots(1)$$

$$\therefore \qquad \mathsf{E}(I_B X) = \int_{-\infty}^\infty I_{X \geq t}\, f(x)\, dx = \lambda \int_t^\infty x\, e^{-\lambda x} dx.$$

Integration by parts provides

$$\mathsf{E}(I_B X) = \lambda \left\{ \frac{x}{\lambda} e^{-\lambda x} + \frac{e^{-\lambda x}}{\lambda^2} \right\}_\infty^t = \lambda e^{-\lambda t} \left(\frac{t}{\lambda} + \frac{1}{\lambda^2} \right)$$

Substitutions into (1) give

$$\mathsf{E}(X \mid X \geq t) = t + (1/\lambda).$$

***Example* 3.** Show that the minimal value of the set $\{X_j \sim \text{expo}(\lambda_j), j = 1, 2, ..., n\}$ is independent of the (ranks) ordering of expo (λ_j).

Solution. Let $j_1, j_2, ..., j_n$ be a permutation of set $\{1, 2, ..., n\}$. Now

$$P\{\min X_j > x \mid X_{j_1} < ... < X_{j_n}\} = \frac{P\{(X_{j_1} < ... < X_{j_n}) \mid \min X_j > x\}\, P\{\min X_j > x\}}{P\{X_{j_1} < ... < X_{j_n}\}} \quad ...(1)$$

where we used Reversal Identity

$$P(A \mid B) = P(B \mid A)\,.\,P(A)/P(B).$$

The memoryless property of exponential variates provides

$$P\{X_{j_1} < ... < X_{j_n} \mid \min X_j > x\} = P\{X_{j_1} < ... < X_{j_n}\} \quad ...(2)$$

Use of (2) into (1) leads to

$$P\{\min X_j > x \mid X_{j_1} < ... < X_{j_n}\} = P\{\min X_j > x\}\,.$$

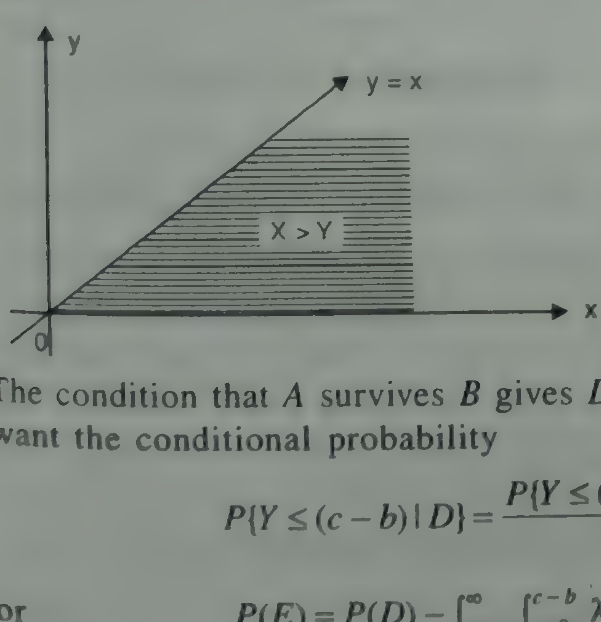
***Example* 4.** Two independent components A and B, each having Expo (λ) distribution are found to be still operating at time t_0 when their ages are a and b. What are the probabilities that (i) both will survive a further time t. (ii) that A will survive B. Given further that A does survive B, find the chance that B's age at failure was at most c ?

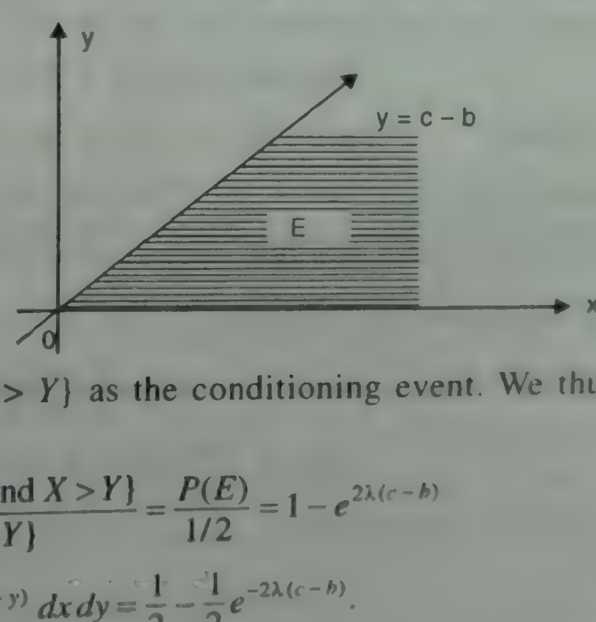
Solution. Since components lack memory, the fact that they are, at time t_0, aged a and b is irrelevant. The future life-times X and Y of A and B have density $\lambda e^{-\lambda x}, (x > 0)$; $\lambda e^{-\lambda y}, (y > 0)$.

Thus (i) and (ii) are essentially unconditional probabilities.

(i) $P\{X > t, Y > t\} = P(X > t)\,.\,P(Y > t) = e^{-2\lambda t}$. $\quad [F(k) = 1 - e^{-\lambda k}]$

(ii) $P(X > Y) = \int_0^\infty \lambda e^{-\lambda y}\, dy\,.\int_y^\infty \lambda e^{-\lambda x}\, dx = \lambda \int_0^\infty e^{-2\lambda y}\, dy = \frac{1}{2}$.





The condition that A survives B gives $D = \{X > Y\}$ as the conditioning event. We thus want the conditional probability

$$P\{Y \le (c-b) \mid D\} = \frac{P\{Y \le (c-b) \text{ and } X > Y\}}{P\{X > Y\}} = \frac{P(E)}{1/2} = 1 - e^{2\lambda(c-b)}$$

or $$P(E) = P(D) - \int_{x=y}^{\infty} \int_{y=0}^{c-b} \lambda^2 e^{-\lambda(x+y)}\, dx\, dy = \frac{1}{2} - \frac{1}{2} e^{-2\lambda(c-b)}.$$

***Example* 5.** At a By-pass, the cars arrive at a Poisson rate of λ per minute and buses arrive at a Poisson rate of μ per minute. The arrivals are independent of each other. Given that the next arrival is a car, find the expected value of the time until the next arrival.

Solution. Let X be the time until the next car arrives at the By-pass and Y be the time until the next bus arrives there. Then $X \sim \text{expo}\left(\frac{1}{\lambda}\right)$ and $Y \sim \text{expo}\left(\frac{1}{\mu}\right)$ are independent.

Let T be the time until the next vehicle arrives at the By-pass and set

$A = \{$Next arrival at the By-pass is a car$\}$.

We need find $E(T \mid A) = E(X \mid X < Y)$. Now

$$P\{X \le t \mid X < Y\} = P\{X \le t, X < Y\}/P\{X < Y\} \qquad \ldots(1)$$

We use conditioning : $P(B) = \sum P(A_j)\, P(B \mid A_j)$; for continuous variates.

$$P\{X < Y\} = \int_0^\infty P[X < Y \mid (X = x)]\, f_X(x)\, dx = \int_0^\infty P(Y > x)\,.\,\lambda\, e^{-\lambda x}\, dx$$

$$= \int_0^\infty \left\{\int_x^\infty \mu\, e^{-\mu y}\, dy\right\} \lambda\, e^{-\lambda x}\, dx = \int_0^\infty \lambda\, e^{-(\lambda+\mu)x}\, dx = \frac{\lambda}{(\lambda+\mu)} \qquad \ldots(2)$$

$$P[X \le t, X < Y] = P[X < \min\{t, Y\}], \text{ use conditioning}$$

$$= \int_0^\infty P[X < \min\{t, Y\} \mid X = x]\, f_X(x)\, dx$$

$$= \int_0^\infty P[\min\{t, Y\} > X \mid X = x]\, f_X(x)\, dx$$

$$= \int_0^t P(Y > x)\,.\, f_X(x)\, dx = \int_0^t \left[\int_x^\infty \mu\, e^{-\mu y}\, dy\right] \lambda\, e^{-\lambda x}\, dx$$

$$= \int_0^t \lambda\, e^{-\mu x}\,.\, e^{-\lambda x}\, dx = \frac{\lambda}{\lambda+\mu}\left[1 - e^{-(\lambda+\mu)t}\right] \qquad \ldots(3)$$

Substitute from (2) and (3) into (1) to get

$$P\{X \le t \mid X < Y\} = 1 - e^{-(\lambda+\mu)t} \qquad \ldots(4)$$

Result (4) shows that, given the next arrival is a car, the distribution of the time until the next arrival is expo $(\lambda + \mu)$. Hence

$$E(T \mid A) = 1/(\lambda + \mu); \qquad \text{[reciprocal of parameter]}$$

Similarly $\qquad E(T \mid \bar{A}) = 1/(\lambda + \mu)$.

***Example* 6.** If X and Y are independent variates with a common Expo (λ) density, find the p.d.f. of $X - Y$ and identify the distribution.

Solution. Recall : $\phi(t) = \int_0^\infty e^{-itz}\, \lambda e^{-\lambda z}\, dz = \dfrac{\lambda}{\lambda - it}$ $\qquad [X \sim \text{Expo}(\lambda)]$

Since X and Y are independent, we have

$$\phi(t : X - Y) = \phi(t : X)\, \phi(-t : Y) = \frac{\lambda}{\lambda - it} \cdot \frac{\lambda}{\lambda + it} = \frac{\lambda^2}{\lambda^2 + t^2}$$

This is well-known Ch. Function of Laplace distribution [§ 8-16(13), $\mu = 0$]

$$f(u) = \tfrac{1}{2}\lambda e^{-\lambda |u|}, (-\infty < u < \infty) \qquad \ldots(1)$$

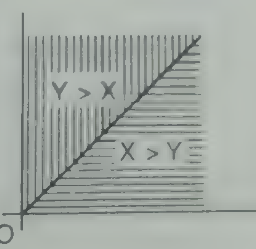
Hence, by Uniqueness. Theorem of Ch. Functions, $U = X - Y$ is Lap $(\lambda, 0)$.

Comment. This problem shows how to obtain Double-exponential distribution (i.e. Laplace) from exponential distribution.

***Example* 7.** Let X and Y be i.i.d. Expo (λ) variates. Show that $Z = \min(X, Y)$ and $D = \max(X, Y) - \min(X, Y)$, are independent.

Solution. Here $z = \min(x, y)$, $d = \max(x, y) - \min(x, y) = |x - y|$; so

$$M(t_1, t_2) = \mathrm{E}(e^{t_1 Z + t_2 D}) = \int_0^\infty \int_0^\infty e^{t_1 z + t_2 d} f(x, y)\, dx\, dy$$



$$= \iint_{x \geq y} e^{t_1 y + t_2 (x - y)} f(x, y)\, dx\, dy + \iint_{x < y} e^{t_1 x + t_2 (y - x)} f(x, y)\, dx\, dy$$

$$= 2 \iint_{x \geq y} e^{t_1 y + t_2 (x - y)} \lambda^2 e^{-\lambda (x + y)}\, dx\, dy. \quad [\text{by symmetry } P(X \geq Y) = P(Y \geq X)]$$

Put $x - y = v$, $y = u$; $|\partial(x, y)/\partial(u, v)| = 1$ so $dxdy = dudv$.

$$\therefore \quad M(t_1, t_2) = \left(\int_0^\infty 2\lambda e^{-2\lambda u} e^{t_1 u} du\right)\left(\int_0^\infty \lambda e^{-\lambda v} e^{t_2 v} dv\right) = \left(1 - \frac{t_1}{2\lambda}\right)^{-1} \cdot \left(1 - \frac{t_2}{\lambda}\right)^{-1} = M(t_1)\, M(t_2).$$

This shows that Z and D are independent; $Z \sim$ Expo (2λ) and $D \sim$ Expo (λ).

Note. Use of *Ordered Statistics* disposes off the problem instantly.

$$\max(x + y) = \tfrac{1}{2}(x + y) + \tfrac{1}{2}|x - y|; \; \min(x + y) = \tfrac{1}{2}(x + y) - \tfrac{1}{2}|x - y|.$$

***Example* 8.** Variates X and Y are jointly distributed as

$$f(x, y) = \lambda^2 x e^{-\lambda x (1 + y)}, \lambda > 0, x > 0, y > 0. \qquad \ldots(1)$$

Show that X and XY are i.i.d. Expo (λ)-variates.

Solution. $U = X$, $V = XY$ and put $x = u$, $y = v/u$. Then

$$\frac{\partial(x, y)}{\partial(u, v)} = \begin{vmatrix} \partial x/\partial u & \partial x/\partial v \\ \partial y/\partial u & \partial y/\partial v \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ -v/u^2 & 1/u \end{vmatrix} = \frac{1}{u}.$$

Thus, $dx\, dy = |1/u|\, dudv$. The joint p.d.f. of U and V is thus, using (1)

$$f(x, y) = dx\, dy = |J|\, g(u, v)\, du\, dv = \lambda^2 u e^{-\lambda u (1 + v/u)} \cdot (1/u)\, du\, dv$$

i.e. $g_1(u, v) = (\lambda e^{-\lambda u})(\lambda e^{-\lambda v}), u > 0, v > 0.$

Thus, U and V are independent Expo (λ)-variates.

***Example* 9.** Show that the exponential variate inherits the memoryless property from the geometric probability law.

Solution. Recall : X is memoryless iff; $P\{X > a + b \mid X > a\} = P\{X > b\}$

i.e. iff : $\quad P\{X > a + b\} = P\{X > a) \cdot P(X > b\}, \qquad \forall\, a, b \geq 0.$

Both geometric and exponential distributions obey this law. [§14-30 and 18-63]

In a sequence of Bernoulli trials, the probability that the first success occurs on trial k is given by gem (p), i.e. $P(X = k) = pq^{k-1}$, $k = 1, 2, \ldots$, so survival function is

$$P(X > n) = \sum_{k = n + 1}^{\infty} pq^{k - 1} = pq^n (1 + q + q^2 + \ldots) = q^n = (1 - p)^n.$$

We now use *Poisson Process* as under : Suppose each trial takes time Δt to perform so that the n trials take $n\, \Delta t$ time. We let $n \to \infty$, $\Delta t \to 0$ in such a way that the total time

consumed remains constant, i.e. $(n\,\Delta t) = t$. (fixed). Now let $p \to 0$ in such a way that the mean number of successes per unit time remains constant. As such we may set $p = \lambda.\Delta t$, where λ is the mean rate of occurrence of the events. Then

$$P(X > n) = (1-p)^n = (1-\lambda\Delta t)^n\,[1-(\lambda t/n)]^n \to e^{-\lambda t},\ \text{as } n \to \infty.$$

Thus, the probability that the first success occurs after time t is $e^{-\lambda t}$. In other words, if T is a variate denoting time until first success, then

$$P(T > t) = e^{-\lambda t} \Rightarrow F(t) = P(T \le t) = 1-(T > t) = 1-e^{-\lambda t} \Rightarrow f(t) = F'(t) = \lambda e^{-\lambda t}, t > 0$$

which is the exponential density to the effect that the first success occurs at time t.

Problems with Solutions Provided at the End of the Text

1*. Prove or disprove : If X and Y are independent, then
(a) mode $(X + Y)$ = mode X + mod Y, (b) med $(X + Y)$ = med X + med Y.

2*. The daily consumption of milk in excess of 20,000 gallons is approximately exponentially distributed with $\lambda = 1/3000$. The city has a daily stock of 35,000 gallons. What is the probability that of two days selected at random, the stock is insufficient for both the days ?

3*. Let $X \sim$ Expo (λ). Find the value x_a such that $P(X > x_a\}/P\{X \le x_a\} = a$.

4*. If $X \sim$ Expo (λ) with $P(X \le 1) = P(X > 1)$, find σ_X^2.

5*. Show that $Y = -(1/\lambda)\ \ln F(X)$ is Expo (λ).

6*. Let X and Y be i.i.d. Expo (λ) distributed. Find the waiting time distribution of $T = Y/X$.

7*. Supoose that the time until failure of some device has an exponential distribution with mean life-time of n months. If k independent devices are observed, find the chance that the first failure occurs within m ($< n$) months.

8*. The time interval T between consecutive random events is Expo (λ).
(a) Show that the expected length of an interval is $1/\lambda$ and find the probability that an interval exceeds this value.
(b) Find the prob. that, of five independent intervals, exactly two are less than the expected value.

9*. Independent r.v. X and Y have exponential p.d.f. $f(x) = \alpha e^{-\alpha x}$, $g(y) = \beta e^{-\beta y}$ $0 < x, y < \infty$.
Find the p.d.f. of $X - Y$.

10*. Find the transformation $Y = H(X)$ such that if X is Expo (λ) then Y has the p.d.f.

$$g(y) = \tfrac{1}{2}\,y^{-1/2},\ 0 \le y \le 1;\ g(y) = 0,\ \text{elsewhere}.$$

11*. Let X be a mixed variate such that

$$f(x) = \begin{cases} \tfrac{1}{2}\lambda^{-1}\,e^{-x/\lambda}, & x > 0, x \text{ not integer} \\ 0, & \text{otherwise} \end{cases} \qquad p(x) = \frac{e^{-\lambda}\lambda^x}{2(x\,!)},\ x = 0, 1, 2, \ldots$$

Find the m.g.f. for the variate X. Hence or otherwise find Var (X).

12*. Independent variates X, Y_1, Y_2, ..., Y_N; N are such that
$P\{X = m\} = (e-1)e^{-m}$; $\quad P\{N = n\} = (e-1)^{-1}/n\,!$, $m, n = 1, 2, \ldots$;
$Y_j \sim U(0, 1)$. Show that $Z = X - \max\{Y_1, Y_2, \ldots, Y_N\} \sim$ Expo (1).

13*. In a sample of n random observations form Expo (λ) distribution, the number of observations in $(0, 1/\lambda)$ and $(1/\lambda, 2/\lambda)$ denoted by X and Y are noted. Find Corr (X, Y).

14*. Uncorrelated variates X and Y are jointly distributed with marginals $X \sim$ Expo(a) and $Y \sim$ Expo(b). Determine the joint distribution.

Exercise 18(b)

1. If X has exponential distribution with mean λ, find $P\{X < 1 \mid X < 2\}$. [**Ans.** $\sqrt{e}/(1+\sqrt{e})$]
2. If $X \sim$ Expo (λ), find p.d.f. of $Y = \log X$ and $X = \log Z$.
   [**Ans.** $f(y) = \lambda \exp(y - \lambda e^y)$; $g(z) = \lambda/z^{1+\lambda}, z \geq 1$]
3. The length of time (in minutes) that a certain lady speaks on the telephone is a random variable specified by p.d.f. as $f(x) = Ae^{-x/5}, x \geq 0$; $f(x) = 0$, otherwise.
   (a) Evaluate A. What is the probability that the number of minutes she takes on phone is (i) more than 10, (ii) less than 5, (iii) between 5 and 10 ?
   (b) Show that, for any two pos. numbers a and b; $P\{X > a + b \mid X > a\} = P\{X > b\}$.
   [**Ans.** $A = 1/5, e^{-2}, 1 - e^{-1}, (e-1)/e^2$]
4. $X \sim U(0, 2)$ and $Y \sim$ Expo(λ). Find the value of λ such that $P(X < 1) = P(Y < 1)$.
5. $X \sim U(-1, 3)$ and $Y \sim$ Expo(λ). Find λ such that $\sigma_X^2 = \sigma_Y^2$. [**Ans.** $\lambda = \sqrt{3}/2$]
6. $X \sim$ gem (p) and $Y \sim$ Expo(λ). Find λ such that $P(X > 1) = P(Y > 1)$. [$\lambda = \ln(1-p)^{-1}$]
7. If $X \sim$ Expo (λ) evaluate $P\{\mid X - \mu \mid \geq k\sigma\}$ and compare it with Chebyshev's bound.
   [**Ans.** $1 - [(e^{2k} - 1)/e^{k+1}]$, bd $1/k^2$]
8. Suppose X is Expo (λ), and define Y by $Y = n$, iff $n - 1 \leq X \leq n, n = 1, 2, 3, \ldots$ Show that Y is gem (p) and find p. [**Ans.** $p = 1 - e^{-\lambda}$]
9. Find the Ch. Function of $Y = -\ln F(X)$, where $F(X)$ is the c.d.f of X. Also evaluate the rth moment of Y. [**Ans.** $\mu'_r = r\,!$]
10. If $F_X(x) = 2\lambda x e^{-\lambda x^2}, x \geq 0$; show that $Y = X^2$ is Expo (λ).
11. If X is Expo (λ), find the distribution of $Y = X/(1 + X)$.
    [**Ans.** $g(y) = \lambda(1-y)^{-2} \exp[-\lambda y/(1-y)], 0 < y < 1$]
12. If X and Y are i.i.d. Expo (λ) variates, find the p.d.f of $X + Y$, $X - Y$ and E $\mid X - Y \mid$.
    [**Ans.** gam $(2, \lambda)$, Lap $(0, \lambda)$, E $\mid Z \mid = \lambda^{-1}$]
13. If X and Y are i.i.d. Expo (λ) variates, find the p.d.f of
    (a) X^3 (b) $2X + 3$ (c) $X^{1/k}$ (d) $X - Y$ (e) $\mid X - Y \mid$.
    [**Ans.** $f_1 = (\lambda/3)w^{-2/3} \exp(-\lambda w^{1/3}), w > 0$, $f_2 = (\lambda/2) \exp[-\lambda(w-3)/2], w > 3$;
    $f_3 = (k/\lambda)w^{k-1} \exp(-\lambda w^k), w > 0$, f_4 See (Q.12), $f_5 = \lambda e^{-\lambda t}, t > 0$]
14. If X be Expo (1), find the p.d.f. of Y where $Y = X$, if $X \leq 1$, $Y = (1/X)$, if $X > 1$.
15. If X, Y, Z are i.i.d. Expo (λ) variates, find $P\{X \geq 2Y \geq 3Z\}$. [**Ans.** 2/33]
16. If $X \sim U(0, 1)$ and $Y \sim$ Expo (λ) are indep. variates, find the p.d.f of (a) Y/λ, (b) XY.
17. If $X \sim$ Expo (a) and $Y \sim$ Expo (b) are indep. variates, find the p.d.f. of Z where $Z = X - Y$, if $X \geq Y$; $Z = 0$, if $X < Y$.

18. If X_k, $1 \le k \le n$, are i.i.d. Expo (λ) variates, find the p.d.f. of $S = X_1 + X_2 + ... + X_n$. If $M_n = \max(X_1, ..., X_n)$ show that $E(M_n) - E(M_{n-1}) = 1/n\lambda$. Hence find $E(M_n)$.

[**Ans.** $(1/\lambda)[1 + (1/2) + (1/3) + ... + (1/n)]$

19. If X is N-B $(1, p)$ and Y is Expo (λ) show that $P\{X \le x\} = P\{Y \le [x]\}, x > 0$, where $\lambda = -\ln(1-p)$; $[x]$ is the greatest integer less than x.

20. Let $X_1, X_2, ...$ be a sequence of i.i.d. Expo (λ) variates. Let $Y = X_1 + X_2 + ... + X_N$, where N is gem (p). Prove that Y is exponentially distributed.

21. Let $Z = \min(X, Y)$, where $X \sim$ Expo (λ) and $Y \sim$ Expo(μ) are independent. Show that

(i) $E\{ZX \mid Z = X) = 2/(\lambda + \mu)^2$,

(ii) $E(ZX \mid Z = Y) = \{2/(\lambda + \mu)^2\} + (1/\lambda, (\lambda + \mu))\}$

(iii) $E(XZ) = (2\lambda + \mu)/\lambda(\lambda + \mu)^2$. Determine Cov (X, Z).

22. Suppose X and Y are i.i.d. Expo (λ) variates. Let $U = \min(X, Y)$; $V = \max(X, Y)$. Find the joint p.d.f. of U and V and hence show that $U \sim$ Expo (2λ) and $V \sim$ Expo (λ) are independently distributed.

23. Let X and Y be two indep. $N(0, 1)$ variates. Let $Z \sim U(0, 1)$ be independent of X and Y and define $W = ZX + (1 - Z)Y$.

(i) Find $E(W)$ and Var (W), (ii) Find the conditional distribution of W given $Z = z$, (iii) Find the distribution of W.

24. Let X be Expo (1). Let $X = x$, take n independent observations $X_1, X_2, ..., X_n$, each $X_i \sim$ Expo (λ). Find the conditional density of X given $(X_1, X_2, ..., X_n)$.

25. Let $F(x)$ and $f(x)$ be the c.d.f. and p.d.f. of a variate $X \ge 0$. Show that to the first order of magnitude.

$$P\{x < X \le x + dx \mid x < X\} = f(x)\, dx / [1 - F(x)] = \mu(x)\, dx.$$

Establish : $F(x) = 1 - \exp\left(-\int_0^x \mu(y)\, dy\right)$.

Deduce that $\mu(x) = \lambda$ (constant) iff X is Expo (λ).

26. If X is Expo(λ), determine $E[X \mid X < c]$ by (i) Definition of conditional mean, (ii) Multistage E-Rule.

27. Records of an insurance company reveal that :

(i) There is a time lag between the reporting of a claim and its settlement :

(ii) The frequency density function for the number of claims getting settled, out of an initial number N of claims, at a point of time in the vicinity of time t (in years) is Nke^{-kt}, $k > 0$.

(iii) The amount paid out as claim increases with the time lag between the reporting and settlement of a claim, the average amount paid out in respect of claims settled at time t being e^{ct}, $c > 0$.

(a) Show that the proportion of claims outstanding at time t is e^{-kt}.

(b) Find the average amount paid in respect of all claims.

(c) Show that the average amount paid in respect of the claims settled within the first year of its being reported is $k(e^k - e^c)/(k - c)(e^k - 1)$.

How does this compare with the result of (b) ?

Fear is the tax that the conscience pays to guilt. (George Sewell)

Try to walk through life instead of being dragged through it.

More Continuous Distributions 19

GAMMA DISTRIBUTION

19-10. *Definition*

The continuous variate X which has p.d.f.

$$f(x) = \frac{e^{-\lambda x}\, x^{\alpha-1}}{\Gamma(a)/\lambda^{\alpha}},\ \lambda > 0,\ a > 0,\ 0 \le x < \infty\ ;\ \ f(x) = 0,\ \text{otherwise}$$

is called a (general) gamma variate with parameters α and λ ; f is called a gamma density and if X has this density we write "X is gam (α ; λ) or $\Gamma(\alpha\,;\lambda)$. The special case $\lambda = 1$ is the *simple* (or *standard*) gamma density and often written gam (α). When $\alpha = n$ is an integer ; gam (n, λ) is called Erlang distribution.

Comments. Another equivalent definition is cited as

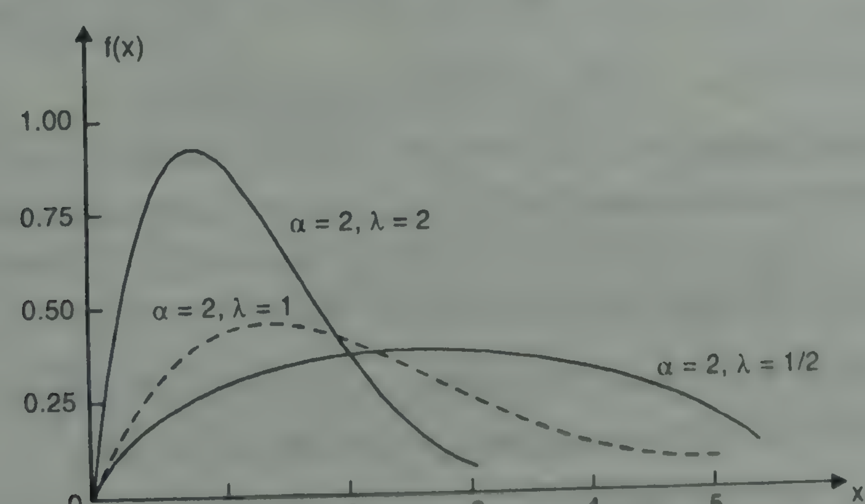
$$f(x) = x^{\alpha-1}\, e^{-x/\beta} / \Gamma(\alpha) . \beta^{\alpha},\ \ \alpha > 0,\ \beta > 0,\ \ x > 0.$$

It is notated X ~ gam (α, β), although $\beta = 1/\lambda$. Be careful about notations.

Cumulative Distribution Function (c.d.f.)

$$F(x) = \int_0^x f(t)\, dt\ \ t > 0;\ \ F(x) = 0,\ \text{otherwise}$$

The c.d.f. (when $\lambda = 1$) is called **incomplete gamma function.**



GAMMA DISTRIBUTION FOR SOME VALUES OF (α, λ)

19-11. Simple Moments

$$E(X^r) = \int_0^\infty \frac{\lambda^\alpha e^{-\lambda x} x^{\alpha-1}}{\Gamma(\alpha)} x^r\,dx = \frac{\lambda^\alpha}{\Gamma(\alpha)}\int_0^\infty e^{-\lambda x} x^{\alpha+r-1}\,dx = \frac{\lambda^\alpha}{\Gamma(\alpha)}\,\frac{\Gamma(\alpha+r)}{(\lambda)^{\alpha+r}}$$

$$= \frac{\Gamma(\alpha+r)}{\lambda^r\,\Gamma(\alpha)} = \frac{\alpha^{[r]}}{\lambda^r}. \qquad [\alpha^{[r]} = \alpha(\alpha+1)\ldots(\alpha+r-1)]$$

This provides μ'_1 = mean = α/λ, $\mu'_2 = \alpha(\alpha+1)/\lambda^2$ and hence Var $(X) = \mu'_2 - \mu'^2_1 = \alpha/\lambda^2$. It follows that for a gam (α, λ) distribution Mean $<, =, >$ Variance $\Leftrightarrow \lambda >, =, < 1$.

Note. For gam (α, β), $E(X) = \alpha\beta$ = product of parameters ; Var $(X) = \alpha\beta^2$.

19-12. M.G.F, Ch. Function, Cumulants and Reproductive Property

$$M(t:X) = [1-(t/\lambda)]^{-\alpha}\ ;\ M(it:X) = [1-(it/\lambda)]^{-\alpha} \qquad [\S\, 8-16\,(9)]$$

$$K(t:X) = \ell n\, M(t) = \ell n[1-(t/\lambda)]^{-\alpha} = \alpha\sum_{r=1}^{\infty}\frac{(t/\lambda)^r}{r} = \alpha\sum_{r=1}^{\infty}\frac{(r-1)!}{\lambda^r}\cdot\frac{t^r}{r!}.$$

Thus, $k_r = \alpha(r-1)!/\lambda^r$. [Coeff. of $t^r/r!$ in $K(t:X)$]

So, $k_1 = \mu'_1 = \alpha/\lambda,\ k_2 = \mu_2 = \alpha/\lambda^2,\ k_3 = \mu_3 = 2\alpha/\lambda^3,\ k_4 = 6\alpha/\lambda^4, \mu_4 = k_4 + 3k_2^2 = (6\alpha + 3\alpha^2)/\lambda^4$.

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{4}{\alpha},\quad \beta_2 = \frac{\mu_4}{\mu_2^2} = 3 + \frac{6}{\alpha},\qquad \gamma_1 = \frac{2}{\sqrt{\alpha}},\ \gamma_2 = \frac{6}{\alpha}.$$

Since $\gamma_1 > 0$, the gamma distribution is positively skewed, and $\gamma_2 > 0$ implies that it is Leptocurtic. Obviously, the point (β_1, β_2) lies on the st. line $2\beta_2 - 3\beta_1 - 6 = 0$.

(ii) If X_k, $1 \le k \le n$ are indep. gam $(\lambda\,;\alpha_k)$ variates, then $(X_1 + X_2 + \ldots + X_n) \sim$ gam $(\lambda\,; \Sigma\alpha_i)$.

Proof. Since $X_i \sim \text{gam}(\lambda, \alpha_i)$, $\quad M(t:X_i) = (1-t/\lambda)^{-\alpha_i}$; so

$M(t:S_n) = M(t:X_1 + X_2 + \ldots + X_n) = M(t:X_1).M(t:X_2)\ldots M(t:X_n)$ [X_i are indep. variates]

$$= (1-t/\lambda)^{-\alpha_1}.(1-t/\lambda)^{-\alpha_2}\ldots(1-t/\lambda)^{-\alpha_n} = [1-(t/\lambda)]^{-\alpha}, \qquad (\alpha = \Sigma\alpha_i)$$

Since m.g.f. uniquely identifies a distribution, it follows that $S_n \sim$ gam $(\lambda : \Sigma\alpha_i)$.

Sum of Exponential Variates : If $X_1, X_2, \ldots, X_n$ are i.i.d. Expo (λ) variates, then

$S_n = X_1 + X_2 + \ldots + X_n$ is a gamma (n, λ) variate.

Proof. For $X \sim$ Expo (λ), $M(t:X) = [1-(t/\lambda)]^{-1}$. Now

$$M(t:S_n) = M(t:X_1 + X_2 + \ldots + X_n) = M(t:X_1)\,M(t:X_2)\ldots M(t:X_n) = [1-(t/\lambda)]^{-n}$$

as X_i are i.i.d. Expo (λ). This shows that $S_n \sim$ gam (n, λ) variate.

19-13. Limiting Distribution

Let $X \sim$ gam (α, λ). Then as $\alpha \to \infty$ (while λ remains constant), $X^* \to N(0, 1)$.

Proof. For gam (α, λ), $\mu = \alpha/\lambda$, $\sigma^2 = \alpha/\lambda^2$, $M(t:X) = [1-(t/\lambda)]^{-\alpha}$. Now let

$$Z = \frac{X-(\alpha/\lambda)}{\sqrt{\alpha}/\lambda} = \frac{\lambda X}{\sqrt{\alpha}} - \sqrt{\alpha} \qquad \text{[Normalized gamma variate]}$$

$$M(t:Z) = M[t:\lambda(\alpha)^{-1/2}X-\sqrt{\alpha}] = e^{-t\sqrt{\alpha}}M(\lambda t\alpha^{-1/2}:X) = e^{-t\sqrt{\alpha}}[1-(t/\sqrt{\alpha})]^{-\alpha}$$

$$\ln M(t:Z) = -t\sqrt{\alpha} - \alpha\ln[1-(t/\sqrt{\alpha})] = -t\sqrt{\alpha} + \alpha[(t/\sqrt{\alpha})+(t^2/2\alpha)+\ldots] = \tfrac{1}{2}t^2 + O(1/\sqrt{\alpha})$$

$\lim_{\alpha\to\infty} \ln M(t:Z) = \frac{1}{2}t^2 \Rightarrow M(t:Z) = e^{t^2/2}$ [$\ln$ function is continuous]. Thus Z is $N(0, 1)$.

19-14. *Mean and Variance of a Truncated Gamma Distribution*

If X is gam (λ, n) i.e. $f(x) = \lambda^n e^{-\lambda x} x^{n-1}/\Gamma(n),\ n>0, \lambda>0, x\ge 0,$ we truncate it and find its probability density in $a \le X \le b$. Recall : Incomplete gamma function

$$\gamma(n,\theta) = \int_0^\theta \frac{e^{-t}t^{n-1}}{\Gamma(n)}dt \qquad \left[\gamma(n+1,\theta) = \gamma(n,\theta) - \frac{e^{-\theta}\theta^n}{\Gamma(n+1)}\right] \qquad \ldots(1)$$

$$\therefore\ \int_a^b f(x)\,dx = \frac{1}{\Gamma(n)}\int_{\lambda a}^{\lambda b} e^{-y}y^{n-1}dy = \frac{1}{\Gamma(n)}\left[\int_0^{\lambda b} - \int_0^{\lambda a}\right]e^{-y}y^{n-1}dy = \gamma(n,\lambda b) - \gamma(n,\lambda a).$$

Thus, the p.d.f. of truncated gam (n, λ)

$$g(x) = A\lambda^n e^{-\lambda x}x^{n-1}/\Gamma(n),\ a\le x\le b. \qquad \{A = [\gamma(n,\lambda b)-\gamma(n,\lambda a)]^{-1}\}$$

$$\therefore\quad \mathsf{E}(X^r) = \int_a^b x^r g(x)\,dx = \frac{\Gamma(n+r)}{\Gamma(n)}\frac{A}{\lambda^r}\int_{\lambda a}^{\lambda b}\frac{e^{-y}y^{n+r-1}dy}{\Gamma(n+r)}$$

$$= \frac{\Gamma(n+r)}{\lambda^r\Gamma(n)}\cdot\frac{\gamma(n+r,\lambda b)-\gamma(n+r),\lambda a)}{\gamma(n,\lambda b)-\gamma(n,\lambda a)} \qquad \ldots(2)$$

$$\therefore\ \mathsf{E}(X) = \frac{n}{\lambda}\cdot\frac{\gamma(n+1,\lambda b)-\lambda(n+1,\lambda a)}{\gamma(n,\lambda b)-\gamma(n,\lambda a)},\ \mathsf{E}(X^2) = \frac{n(n+1)}{\lambda^2}\cdot\frac{\gamma(n+2,\lambda b)-\lambda(n+2,\lambda a)}{\gamma(n,\lambda b)-\gamma(n,\lambda a)} \qquad \ldots(3)$$

So Var $(X) = \mathsf{E}(X^2) - \mathsf{E}^2(X)$, is instantly obtained. The incomplete gamma function in (3) can be reduced by reduction formula [1 (ii)]. However, the final result is not elegant.

19-15. Worked-out Problems

***Example* 1.** Prove that : $\displaystyle\sum_{j=0}^{n}\frac{e^{-\lambda}\lambda^j}{j!} = \frac{1}{n!}\int_\lambda^\infty e^{-x}x^n\,dx.$ ...(1)

Hence write down the relation connecting the distribution function of a Poisson and a gamma or χ^2-variate.

Solution. The series involved is really $F(n) = P\{X\le n\}$ where $X\sim$ Pois (λ). Now

$$\frac{d}{d\lambda}P\{X\le n\} = \frac{d}{d\lambda}\sum_{j=0}^{n}\frac{e^{-\lambda}\lambda^j}{j!} = \sum_{j=0}^{n}\frac{1}{j!}e^{-\lambda}(j\lambda^{j-1}-\lambda^j) = e^{-\lambda}\sum_{j=0}^{n}\left[\frac{\lambda^{j-1}}{(j-1)!}-\frac{\lambda^j}{j!}\right]$$

The extreme member is a Telescopic series and boils down to $-e^{-\lambda}\lambda^n/n!$. Thus

$$DP\{X\le n\} = -e^{-\lambda}\lambda^n/n!,\ [D\equiv d/d\lambda]$$

Integrating this result between 0 and λ, noting $F(n) = 1$ when $\lambda = 0$, we get

$$F(n) - 1 = -\int_0^\lambda \frac{e^{-t}\, t^n}{n!}\, dt \;\Rightarrow\; F(n) = 1 - \int_0^\lambda \frac{e^{-t}\, t^n}{n!}\, dt = \int_\lambda^\infty \frac{e^{-t}\, t^n}{n!}\, dt. \qquad \text{...(2)}$$

This is the result sought. Note that, during integration, we had changed dummy λ to dummy t.

Alternatively, this result can be written as $P\{X \le n\} = P\{Y \ge \lambda\}$ where $Y \sim$ gam $(n + 1, 1)$. It follows that Poisson distribution can be evaluated from the incomplete gamma function and vice-versa.

***Example* 2.** If $X \sim$ gam (r, λ) where r is a positive integer, show that

$$F_X(x) = 1 - \sum_{j=0}^{r-1} \frac{e^{-\lambda x}\, (\lambda x)^j}{j!}.$$

Solution. Evaluations on integration by parts provide :

$$\int_x^\infty t\, e^{-\lambda t}\, dt = \left[(t)\left(-\frac{e^{-\lambda t}}{\lambda}\right) - (1)\left(\frac{e^{-\lambda t}}{\lambda^2}\right)\right]_x^\infty = \frac{1}{\lambda^2}\, e^{-\lambda x}\,(1 + \lambda x)$$

$$\int_x^\infty t^2\, e^{-\lambda t}\, dt = \left[(t^2)\left(-\frac{e^{-\lambda t}}{\lambda}\right) - (2t)\left(\frac{e^{-\lambda t}}{\lambda^2}\right) + 2\left(\frac{-e^{-\lambda t}}{\lambda^3}\right)\right]_x^\infty = \frac{2!}{\lambda^3}\, e^{-\lambda x}\left(1 + \lambda x + \frac{(\lambda x)^2}{2!}\right).$$

$$\int_x^\infty t^3\, e^{-\lambda t}\, dt = \frac{e^{-\lambda x}}{\lambda^4}\,(\lambda^3 x^3 + 3\lambda^2 x^2 + 6\lambda x + 6) = \frac{3!\, e^{-\lambda x}}{\lambda^4}\left(1 + \lambda x + \frac{(\lambda x)^2}{2!} + \frac{(\lambda x)^3}{3!}\right).$$

Thus, by Mathematical Induction

$$\int_x^\infty t^k\, e^{-\lambda t}\, dt = \frac{k!\, e^{-\lambda x}}{\lambda^{k+1}}\left[1 + \lambda x + \frac{(\lambda x)^2}{2!} + \ldots + \frac{(\lambda x)^k}{k!}\right]. \qquad [\text{Take } k = r - 1]$$

So $\displaystyle\int_x^\infty \frac{\lambda^r\, e^{-\lambda t}\, t^{r-1}}{\Gamma(r)}\, dt = \sum_{j=0}^{r-1} \frac{e^{-\lambda x}\, .(\lambda x)^j}{j!} \Rightarrow 1 - F_X(x) = F_Y(r-1)$, or $F_X(x) = 1 - F_Y(r-1)$.

where $X \sim$ gam (r, λ) and $Y \sim$ Pois (λx). The stated result is established.

***Example* 3.** Using C.L.T. show that

$$\lim_{n\to\infty} e^{-n} \sum_{k=0}^{n} \frac{n^k}{k!} = \frac{1}{2} = \lim_{n\to\infty} \int_0^n \frac{e^{-t}\, t^{n-1}}{(n-1)!}\, dt.$$

Solution. (i) Let $X_1, X_2, \ldots, X_n$ be a random sample from Pois (1) and put $S_n = X_1 + X_2 + \ldots + X_n$. Then $S_n \sim$ Pois (n), so that $\mathrm{E}(S_n) = n = \mathrm{Var}\,(S_n)$.

Consequently, $S_n^* = (S_n - n)/\sqrt{n}$. By C.L.T., as $n \to \infty$, $Z = S_n^*$ tends to $N(0, 1)$ distributed.

Now $\qquad P(S_n = k) = e^{-n} n^k / k!$, $[S_n \sim \text{Pois}\,(n)]$ so that

$$P\{S_n \le n\} = \sum_{k=0}^{n} P(S_n = k) = e^{-n} \sum_{k=0}^{n} (n^k / k!) \qquad \text{...(1)}$$

Also $\quad P\{S_n \le n\} = P\left\{\frac{S_n - n}{\sqrt{n}} \le 0\right\} = P\{S_n^* \le 0\}$ and $S_n^* \sim N(0,1)$ for large n. Hence

$$\lim_{n\to\infty} P\{S_n^* \le 0\} = \int_{-\infty}^{0} f(t)\,dt = \frac{1}{2} \qquad f(t) : \text{Normal density.} \qquad \ldots(2)$$

Thus, $$\lim_{n\to\infty} P\{S_n \le n\} \equiv \lim_{n\to\infty} \sum_{k=0}^{n} \frac{e^{-n} n^k}{k!} = \frac{1}{2} \qquad \text{[by (1) and (2)]} \qquad \ldots(3)$$

(ii) Let $X_1, X_2, \ldots, X_n$ be a random sample from Expo (1) and put $T_n = X_1 + X_2 + \ldots + X_n$. Then $T_n \sim$ gam $(n, 1)$ by Reproductive property. $E(T_n) = n = \text{Var}(T_n)$. By C.L.T.

$$\lim_{n\to\infty} P\left\{\frac{T_n - n}{\sqrt{n}} \le x\right\} = \int_{-\infty}^{x} f(t)\,dt \qquad \text{[choose } x = 0]$$

$$\lim_{n\to\infty} P\{T_n \le n\} \equiv \lim_{n\to\infty} \int_0^n \frac{e^{-t} t^{n-1}}{\Gamma(n)}\,dt = \int_{-\infty}^{0} f(t)\,dt = \frac{1}{2} \Rightarrow \lim_{n\to\infty} \int_0^n \frac{e^{-t} t^{n-1}}{(n-1)!}\,dt = \frac{1}{2}. \qquad \ldots(4)$$

We combine (3) and (4) to get the stated result.

Problems with Solutions Provided at the end of the Text

1*. Consumer demand for milk in a certain locality, per month, is known to be a gamma variate. If the average demand is a litres and the most likely demand is b litres ($b < a$), what is the variance of the demand ?

2*. Find the p.d.f. of X if the random variable X is such that

$$E(X^n) = (n+k)!\,k^n/k!, \quad n = 1, 2, 3, \ldots \text{[} k \text{ is pos. integer]}$$

3*. Show that if X is $N(\mu, \sigma^2)$, then $Y = (X - \mu)^2/2\lambda\sigma^2$ is gam $(1/2, \lambda)$. Hence find Var (Y).

4*. Show that the mean value of positive square root of a gam (r, λ), variate is $\Gamma(r + \frac{1}{2})/\sqrt{\lambda}\,\Gamma(r)$. Hence prove that the mean deviation of a $N(\mu, \sigma^2)$ variate from its mean is $\sigma\sqrt{2/\pi}$.

5*. Let X_1, X_2 be a random sample of size 2 from Expo (λ) distribution. Show that $Y_1 = (X_1 + X_2)$ and $Y_2 = X_1/(X_1 + X_2)$ are independent.

6*. A random sample of size n is taken from a gam $(m, 1/a)$ distribution. Find the distribution of $\overline{X}$.

7*. A random sample of size n is taken from Expo (λ) population. If $\overline{X}$ is the sample mean, show that $n\lambda\overline{X}$ is gam $(n, 1)$ and that S.E. of $\overline{X}$ is $1/\lambda\sqrt{n}$.

8*. Derive the m.g.f. of *shifted* (i.e. *truncated*) gamma distribution whose p.d.f. is

$$f(x) = \frac{e^{-\lambda(x-\theta)}\,(x-\theta)^{\alpha-1}}{\Gamma(\alpha)/(\lambda)^{\alpha}}, \quad x \ge \theta, \ \lambda > 0, \ \alpha > 0.$$

If $X_1, X_2, \ldots, X_n$ is a random sample of size n from a shifted exponential distribution $[\alpha = 1]$, find the sampling distribution of $\overline{X}$.

9*. If X is gam (α, λ), show that $P\{X \geq 2\alpha/\lambda\} \leq (2/e)^{\alpha}$.

Exercise 19(a)

1. Let X be gam (α, λ). Show that mode $X = (\alpha - 1)/\lambda$, if $\alpha > 1$ and mode $X = 0$, if $\alpha < 1$.
2. Show that for gam (α, λ) distribution

$$\frac{\text{Mean} - \text{Mode}}{\text{S.D.}} = \frac{1}{\sqrt{\alpha}} = \frac{1}{2}\frac{\mu_3}{\sigma_3}; \text{ Excess of kurtosis} = \frac{6}{\alpha}.$$

3. (a) If X and Y are independent $N(0, 1)$ variates, find the p.d.f. of (i) X^2, (ii) $X^2 + Y^2$, (iii) X^2/Y^2.
   (b) If $X \sim$ gam $(1 + \lambda, \lambda)$, show that $P[0 < X < 2(1 + \lambda)] > \lambda/(1 + \lambda)$.
4. Let X have a p.d.f. $f(x) = \alpha x^2 e^{-bx^2}$, $x > 0$; $f(x) = 0$ elsewhere. Show that $T = \frac{1}{2} mX^2$ is gamma variate. Hence show that $2bT/m$ is gam (3/2).
5. Let $f(x, \theta, r) = \sum_{k=0}^{r-1} \dfrac{[(1 + k\theta)\, x]^{k-1}}{k!} \cdot e^{-(1+k\theta)x}[(1 + k\theta)\, x - k]$, $x > 0$, $r = 1, 2, \ldots, \theta > 1/r$.

   Show that $f(x; \theta, r)$ is a p.d.f. of some variate X. [called Lagrangian gamma density]. Find Var (X) and $M(t : X)$. Show further that $E(X^{-1}) = \ln(1 + \theta) + (r - 1)^{-1}$, $r \geq 2$.
   $$E(X^{-2}) = (1 + 2\theta)\{\ln(1 + 2\theta) - \ln(1 + \theta)\} + (r - 2)^{-1} - E(X^{-1});\ r \geq 3.$$
6. If the conditional distribution of Y given $X = x$ is Expo (x) and the unconditional distribution of X is gamma with parameters $a > 2$, $\lambda > 0$, find the marginal density of Y and its mean and variance. Show that the conditional density of X given $Y = y$ is gam $(\alpha + 1, \lambda + y)$.
7. A two-dimensional random variable (X, Y) has the joint p.d.f.

   $$f(x, y) = [\Gamma(l)\,\Gamma(m)]^{-1} x^{l-1} (y - x)^{m-1} e^{-y}, 0 < x < y < \infty; \quad f(x, y) = 0 \text{ elsewhere.}$$

   Show that the marginal distributions of X and Y are gamma distribution. Find the distribution of $Z = Y - X$. Show that Z and X are independent.
8. If X and Y are independent gam (α, λ) and gam $(\frac{1}{2} + \alpha, \mu)$ variates, show that the variate $Z = 2\sqrt{XY}$ has gam $(2\alpha, \sqrt{\mu\lambda})$ distribution.
9. A random sample $X_1, \ldots, X_n$ of size n is drawn from gam (α, λ) distribution. If $\overline{X}$ is the mean of the sample, find the distribution of $n\overline{X}$. Hence or otherwise find its variance.
10. Let X_i, $1 \leq i \leq n$, be a random sample from a gamma distribution. Show that T/S and S are independently distributed, where $S = \Sigma X_i$, $T = \Sigma a_i X_i$ and a_i's are real numbers.
11. Let X be a non-negative continuous variate, and let Y be $U(0, X)$. If $Z = X - Y$, then Y and Z are independent if X is gam $(\alpha, 2)$ for same $\alpha > 0$.
12. The distribution of a variate X^2 is gam (a) if the density of X is $g(x) = h(x)\,|x|^{2m-1}.\exp(-ax^2)$, where $h(x) + h(-x) = \text{const.}\ \forall\, x$.
13. Let $X \sim$ gam (α, λ). Find Corr (X, Y) where $Y + X^n$.

BETA DISTRIBUTION OF THE FIRST KIND

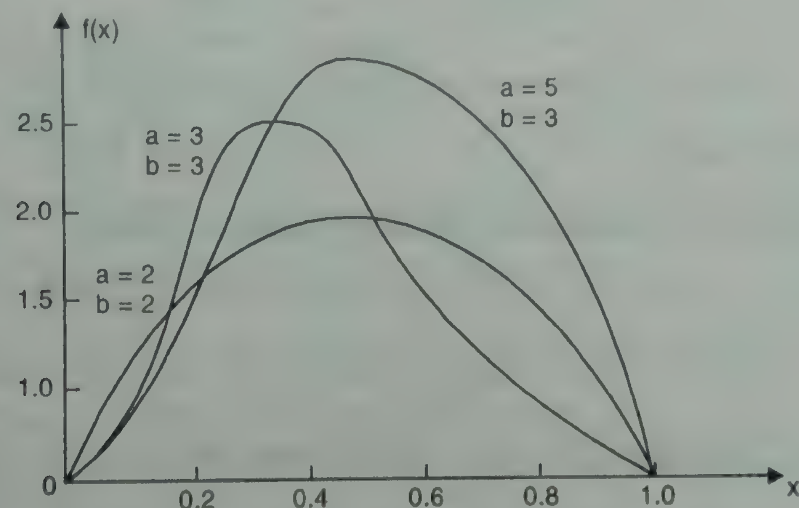
19-20. *Definition*

The continuous r.v. X possessing the p.d.f.

$$f(x) = \frac{x^{a-1}(1-x)^{b-1}}{B(a,b)}, a, b > 0,\ 0 < x < 1; \quad f(x) = 0, \text{ otherwise}$$

is called a beta-I variate with parameters a and b; f is called *beta density* and if X has this density, we write it as $X \sim \beta_{\mathrm{I}}(a, b)$.

Remarks. (i) If $a = b = 1$, $f(x) = 1$ so $X \sim$ Unif $(0, 1)$



(ii) If $a = 2$, $b = 1$ or $a = 1$, $b = 2$, the triangular distribution results :

$$f(x) = \begin{cases} 2(1-x), & 0 < x < 1 \\ 0, & \text{otherwise} \end{cases}; \quad f(x) = \begin{cases} 2x, & 0 < x < 1 \\ 0, & \text{otherwise} \end{cases}$$

Distribution Function :

$$F(x) = \int_0^x \frac{t^{a-1}(1-t)^{b-1}}{B(a,b)}\, dt,\ a, b > 0, 0 < x < 1; F(x) = 0, \text{ if } x < 0\ F(x) = 1, \text{ if } x > 1.$$

The c.d.f. is also known as "Incomplete beta function".

19-21. *Simple Moments and Moments Generating Function*

$$\mu_n' = \mathrm{E}(X^n) = \int_0^1 \frac{x^{a-1}(1-x)^{b-1}}{B(a,b)} x^n\, dx = \frac{1}{B(a,b)} \int_0^1 x^{a+n-1}(1-x)^{b-1}\, dx.$$

i.e. $$\boxed{\mu_n' = \frac{B(a+n,b)}{B(a,b)} = \frac{\Gamma(a+n)}{\Gamma(a)} \cdot \frac{\Gamma(a+b)}{\Gamma(a+b+n)} = \frac{a^{[n]}}{(a+b)^{[n]}}, k^{[n]} = k(k+1)\ldots(k+n-1)} \quad \ldots(1)$$

$$\therefore \quad \mu_1' = \frac{a}{a+b} = \frac{a}{c}\,[c = a+b];\ \mu_2' = \frac{a(a+1)}{c(c+1)};\ \mu_3' = \frac{a(a+1)(a+2)}{c(c+1)(c+2)}$$

$$\mu_4' = \frac{a(a+1)(a+2)(a+3)}{c(c+1)(c+2)(c+3)}.\ \mu_2 = \mu_2' - \mu_1'^2 = ab/c^2(c+1),$$

$$\mu_3 = \frac{2ab(b-a)}{c^3(c+1)(c+2)}, \quad \mu_4 = \frac{3ab[2c^2+ab(c-6)]}{c^4(c+1)(c+2)(c+3)}$$

$$(\alpha_3)^2 = \frac{\mu_3^2}{\mu_2^3} = \frac{4(b-a)^2(c+1)}{ab(c+2)^2}; \quad \alpha_4 = \frac{\mu_4}{\mu_2^2} = \frac{3(c+1)(2c^2+abc-6ab)}{ab(c+2)(c+3)}.$$

The harmonic mean H is given by

$$\frac{1}{H} = \mu'_{-1} = \frac{\Gamma(a+b)}{\Gamma(a)} \cdot \frac{\Gamma(a-1)}{\Gamma(a+b-1)} = \frac{a+b-1}{a-1} \Rightarrow H = \frac{a-1}{c-1}.$$

The parameter $c = a + b$ is often called "concentration parameter".

M.G.F. We use Taylor Series expansion for m.g.f. to get

$$M(t) = \sum_0^\infty \frac{t^x}{x!}\mu'_r = 1 + \sum_{x=1}^\infty \frac{\Gamma(a+b)}{\Gamma(a+b+x)} \cdot \frac{\Gamma(a+x)}{\Gamma(a)} \cdot \frac{t^x}{x!}$$

This result is of little use.

19-22. *Mean Absolute Deviation (M.A.D.) about Mean*

$$M = \mathrm{E}|X-\mu| = \sum_{x>\mu} f(x)(x-\mu) + \sum_{x<\mu} f(x)(\mu-x)$$

$$0 = \mathrm{E}(X-\mu) = \sum_{x>\mu} f(x)(x-\mu) + \sum_{x<\mu} f(x)(x-\mu)$$

Subtracting : $M = 2\Sigma f(x)(\mu-x), \ x<\mu$. Substituting for $f(x)$,

$$\frac{1}{2}B(a,b).M = \int_0^\mu (\mu-x)x^{a-1}(1-x)^{b-1}\,dx, \quad \left(\mu = \frac{a}{a+b} = \frac{a}{c}\right) \qquad \text{...(1)}$$

Let $J(t) = \int_0^\mu (t-x)x^{a-1}(1-x)^{b-1}\,dx, \quad \therefore \frac{dJ}{dt} = J'(t) = \int_0^\mu x^{a-1}(1-x)^{b-1}\,dx$ ...(2)

We rewrite (2i) as

$$J = t\int_0^\mu x^{a-1}(1-x)^{b-1}\,dx - \int_0^\mu x^a(1-x)^{b-1}\,dx = tJ' - K, \ \text{say} \qquad \text{...(3)}$$

$K = \int_0^\mu x^a(1-x)^{b-1}\,dx$ [Integ. by Parts, put $\lambda = 1-\mu$] ...(4)

$$= \frac{\mu^a\lambda^b}{b} + \frac{a}{b}\int_0^\mu x^{a-1}(1-x)^{b-1}(1-x)\,dx = -\frac{\mu^a\lambda^b}{b} + \frac{a}{b}[J'-K], \quad \text{[by (2 ii) and (4)]}.$$

Thus, $K = \mu J' - (\mu^a\lambda^b/c)$; substituting for K in (3) gives $(t-\mu)J' = J - A$. $(A = \mu^a\lambda^b/c)$

$\therefore \dfrac{dJ}{J-A} = \dfrac{dt}{t-\mu}$, integration provides $\dfrac{J-A}{t-\mu}$ = constant (independent of t).

If $t \to \mu$, $J(t) \to A$; i.e. $J(\mu) = A = \mu^a\lambda^b/c$. This evaluates (2). Thus from (1) and (2)

$$M = \frac{2J(\mu)}{B(a,b)} = \frac{2\mu^a\lambda^b}{cB(a,b)} = \frac{2a^a\,b^b}{(a+b)^{a+b+1}}\frac{1}{B(a,b)}.$$

19-23. *Modal Value of beta-I Variate*

Let $y = f(x) = [B(a, b)]^{-1} x^{a-1}(1-x)^{b-1}, \quad 0 < x < 1.$ ...(1)

We first note that if $a < 1$, then $f(x) \to \infty$ as $x \to 0+$, so $x = 0$ is modal value. If $b < 1$, then $f(x) \to \infty$ as $x \to 1-$, i.e. $x = 1$ is mode. If both $a < 1$, $b < 1$ hold, then the distribution is bimodal with modes at $x = 0$ and $x = 1$.

If $a = 1$, $b = 1$, then $f(x) = 1$, $0 < x < 1$; each $x \in (0, 1)$ is mode.

If $a = 1$, $b > 1$, then $f(x) = (1-x)^{b-1}/B(1, b) = b(1-x)^{b-1}$, $0 \le x \le 1$.

Obviously, $f'(x) < 0$, so f is decreasing over $[0, 1]$, hence $x = 0$ is the mode.

If $a > 1$, $b = 1$, then $f(x) = x^{a-1}/B(a, 1). = a\,x^{a-1}, \quad 0 \le x \le 1$

Obviously, $f'(x) > 0$, so f is increasing over $[0, 1]$, hence $x = 1$ is the mode.

Now let $a > 1$, $b > 1$, $c = (a + b) > 2$. Take logarithms in (1) to get

$$\ell n\ y = -\ell n\ B(a,b) + (a-1)\ell n\ x + (b-1)\ell n(1-x).$$

We differentiate it twice w.r.t. "x" and obtain

$$y^{-1}y' = (a-1)x^{-1} - (b-1)(1-x)^{-1};$$

$$y^{-1}y'' - y^{-2}y'^2 = -[(a-1)x^{-2} + (b-1)(1-x)^{-2}] \quad (x \ne 0, x \ne 1)$$

Now $y' = 0 \Rightarrow (1-x)(a-1) = (b-1)x \Rightarrow x = (a-1)/(c-2).$ ...(2)

When $y' = 0$, $y'' < 0$, $(a > 1, b > 1)$ hence (2) gives the modal value.

19-24. *Generalized B_I-Distribution*

The continuous random variable X which has p.d.f.

$$f(x) = \frac{(x-m)^{a-1}(M-x)^{b-1}}{(M-m)^{a+b-1}\,B(a,b)}, \quad m \le x \le M,\ a, b > 0; \qquad f(x) = 0, \text{ otherwise}$$

is called the *general* beta distribution of the first kind. It is easy to verify the following :

$$\mu_X = m + \frac{ha}{c}, \quad \sigma_X^2 = \frac{h^2\,ab}{c^2\,(c+1)}, \quad [c = a + b \text{ and } h = M - m].$$

$$\text{Mode} = m + \frac{h(1-a)}{2-c}, \quad \text{Coeff. of Skewness } S_k = \frac{2(b-a)}{c\,(c+2)\,\sigma_X}.$$

We may also observe that : $P\{p \le X \le q\} = \dfrac{1}{B(a,b)} \displaystyle\int_p^q \frac{(x-m)^{a-1}(M-x)^{b-1}}{(M-m)^{c-1}}\,dx.$

Put $y = \dfrac{x-m}{h}$, $u = \dfrac{q-m}{h}$, $v = \dfrac{p-m}{h}$, so $h(1-y) = (M-x)$ to recover

$$P\{p \le X \le q\} = \frac{1}{B(a,b)}\left\{\int_0^u y^{a-1}(1-y)^{b-1}\,dy - \int_0^v y^{a-1}(1-y)^{b-1}\,dy\right\}$$

$$= B_u\,[(a,b) - B_v\,(a,b)]/B(a,b)$$

where $B_\theta\,(a,b) = \int_0^\theta y^{a-1}(1-y)^{b-1}\,dy.$ [Incomplete beta function]

Note. $B\theta\ (a, b)/B(a, b) = P\{X \le \theta\}$ is beta c.d.f. or in-complete beta function Ratio.

19-25. Worked-out Problems

***Example* 1.** Let $X \sim$ gam (a, λ) and $Y \sim$ gam (b, λ) be independent. Show that
(i) $U = X + Y$, $V = X/(X + Y)$ are independent, U is gam $(a + b, \lambda)$ and V is $B_1(a, b)$
(ii) $X = R\sin^2\theta$, $Y = R\cos^2\theta$, $0 < R < \infty$, $0 < \theta < \pi/2$ emply that R and θ are independent.

Solution. (i) Let $u = x + y$, $v = x/(x + y)$ then $x = uv$, $y = u(1 - v)$

$$\therefore \quad \frac{\partial(x, y)}{\partial(u, v)} = \begin{vmatrix} \partial x/\partial u & \partial y/\partial u \\ \partial x/\partial v & \partial y/\partial v \end{vmatrix} = \begin{vmatrix} v & 1-v \\ u & -u \end{vmatrix} = -u. \text{ Thus, } dx\, dy = |\, u\,|\, du\, dv.$$

The region $x \ge 0$, $y \ge 0$ transforms to the region $u \ge 0$, $0 < v < 1$ in the u-v plane.
Since X and Y are independent, their joint probability element is

$$dP_1(x, y) = K\lambda^{a+b} x^{a-1} y^{b-1} e^{-\lambda(x+y)} dx\, dy, \qquad K = [\Gamma(a)\Gamma(b)]^{-1} \qquad (A)$$

In terms of (u, v) this reduces to

$$dP_2(u, v) = \frac{\lambda^{a+b} e^{-\lambda u} u^{a+b-1} du}{\Gamma(a+b)} \cdot \frac{v^{a-1}(1-v)^{b-1} dv}{B(a, b)} = dF_1(u) \,.\, dF_2(v), (0 < u < \infty, 0 < v < 1) \quad \ldots(1)$$

This result shows that U and V are indep., U is gam $\sim (a + b, \lambda)$ and V is $B_1(a, b)$.

(ii) Here $x = r\sin^2\theta$, $y = r\cos^2\theta$, $\quad \partial(x, y)/\partial(r, \theta) = -2r\sin\theta\cos\theta$

$$\therefore \qquad dx\, dy = 2r\sin\theta\cos\theta\, dr\, d\theta\,.$$

The joint probability element (A), when converted to variables r, θ gives

$$dF(r, \theta) \equiv K(\lambda)^{a+b} e^{-r} (r)^{a+b-2} (\sin\theta)^{2a-2} (\cos\theta)^{2b-2} . 2r\sin\theta\cos\theta\, dr\, d\theta$$

$$= \left\{\frac{(\lambda)^{a+b}(r)^{a+b-1} e^{-\lambda r}\, dr}{\Gamma(a+b)}\right\}\left\{\frac{(\sin\theta)^{2a-1}(\cos\theta)^{2b-1}}{2B(a, b)}\, d\theta\right\}, \quad 0 < r < \infty, \quad 0 < \theta, \pi/2.$$

This result shows that R and θ are independent, and $R \sim$ gam $(a + b, \lambda)$ and $\theta \sim B_1(a, b)$.

***Example* 2.** If $X \sim B_1(a, b)$ and $Y \sim$ gam $(a + b, \lambda)$ are independent, then XY is gam (a, λ).

Solution. Let $u = xy$, $v = x$ so that $x = v$, $y = u/v$. Then $dx\, dy = |\, J\,|\, du\, dv = v^{-1}\, du\, dv$.
The region $0 < x < 1$ and $0 < y < \infty$ in the x-y plane transforms to the region $0 < u < \infty$, $0 < v < 1$ in the u-v plane.
Since X and Y are independent their joint probability element is

$$dP = \frac{x^{a-1}(1-x)^{b-1}\, dx}{B(a, b)}, \frac{e^{-\lambda y} \,.\, y^{a+b-1}\, dy}{\Gamma(a+b)/(\lambda)^{a+b}} = \frac{e^{-\lambda u/v} \,.\, u^{a+b-1}(1-v)^{b-1} v^{-(b+1)} \lambda^{a+b}\, du\, dv}{\Gamma(a)\Gamma(b)}$$

We integrate **out** v to get

$$dF(u) = \frac{\lambda^{a+b} u^{a+b-1}\, du}{\Gamma(a)\Gamma(b)} \int_0^1 e^{-\lambda u/v}\left(\frac{1}{v} - 1\right)^{b-1} \frac{dv}{v^2} \qquad \left[\text{Put } z = \frac{1}{v} - 1\right]$$

$$= \frac{\lambda^{a+b} u^{a+b-1} e^{-\lambda u}\, du}{\Gamma(a)\Gamma(b)} \int_0^\infty e^{-\lambda uz} z^{b-1}\, dz = \frac{\lambda^{a+b} u^{a+b-1} e^{-\lambda u}}{\Gamma(a)\Gamma(b)} \cdot \frac{\Gamma(b)}{(\lambda u)^b}\, du$$

i.e. $dF(u) = \{\lambda^a e^{-\lambda u} u^{a-1}/\Gamma(a)\}\, du$, $0 < u < \infty$. This shows that $U = XY \sim$ gam $(\lambda\,;\, a)$.

N.B. Mellin Transform is also applicable.

Problems with Solutions Provided at the End of the Text

1*. The duration (dur) for a construction project for completion is assessed as under : min dur. = 6 days, max dur. = 12 days, expected dur. = 8 days.
The Coeff. of variation of duration is 10%. What is the probability that the construction project is completed within 10 days ?

2*. If X is a B_I variate with mean μ and concentration parameter $c \neq 2$, find its mode m and verify that : $m - \mu = (2\mu - 1)/(c - 2)$.
What is the least upper bound for $|m - \mu|$ if μ varies and c remains fixed ?

3*. If X is $B_I(a, b)$, then $Y = 1 - X$ is $B_I(b, a)$.

4*. Let $X \sim B_I(a, b)$ and $Y = B_I(p, q)$ independent variates. Show that $XY \sim B_I(p, b + q)$ provided $p + q = a$.

BETA DISTRIBUTION OF THE SECOND KIND

19-30. Definition

The continuous variate X possessing the p.d.f.

$$f(x) = \frac{x^{a-1}}{B(a, b)(1 + x)^{a+b}}, \; a > 0, \; b > 0, \; 0 \le x < \infty; \quad f(x) = 0, \text{ otherwise}$$

is called a B_{II} or beta prime variate, with parameters a and b; f is called $B_{II}(a, b)$ density.

Note. The two kinds of beta distributions are essentially the same. These are designed to cater for the ranges $[0, 1]$ and $[0, \infty[$.

Let $z = \dfrac{x}{1 - x}$ i.e. $x = \dfrac{z}{1 + z}$, then

$$\int_0^1 x^{a-1}(1 - x)^{b-1}\,dx = \int_0^\infty \frac{z^{a-1}\,dz}{(1 + z)^{a+b}}.$$

Distribution Function :

$$F(x) = \frac{1}{B(a, b)} \int_0^x \frac{t^{a-1}\,dt}{(1 + t)^{a+b}}, \quad a, b > 0, 0 < x < \infty; \quad F(x) = 0, \text{ otherwise.}$$

19-31. Simple Moments and M.G.F.

$$\mu'_n = E(X^n) = \int_0^\infty \frac{x^{a-1}\,x^n\,dx}{B(a, b)(1 + x)^{a+b}} = \frac{1}{B(a, b)} \int_0^\infty \frac{x^{a+n-1}\,dx}{(1 + x)^{(a+n)+(b-n)}} = \frac{B(a + n, b - n)}{B(a, b)}$$

$$= \Gamma(a + n)\,\Gamma(b - n) / \Gamma(a)\,\Gamma(b) = a^{[n]} / (b - n)^{[n]} \qquad \text{...(1)}$$

Thus $\mu'_1 = \dfrac{a}{b - 1}$ (Mean); $\quad \mu'_2 = \dfrac{a(a + 1)}{(b - 1)(b - 2)}$, $\quad \mu'_3 = \dfrac{a(a + 1)(a + 2)}{(b - 1)(b - 2)(b - 3)}$

$$\mu'_4 = \frac{a(a + 1)(a + 2)(a + 3)}{(b - 1)(b - 2)(b - 3)(b - 4)}; \qquad \mu_2 = \mu'_2 - \mu_1'^2 = \frac{a(a + b - 1)}{(b - 1)^2 (b - 2)}$$

The Harmonic Mean H is given by

$$\frac{1}{H} = \mu'_{-1} = \frac{\Gamma(a-1).\Gamma(b+1)}{\Gamma(a).\Gamma(b)} = \frac{b}{a-1} \Rightarrow H = \frac{a-1}{b}.$$

$$M(t) = \sum_{r=0}^{\infty} \mu'_r . \frac{t^r}{r!} = 1 + \sum_{r=1}^{\infty} \frac{\Gamma(a+r)\Gamma(b-r)}{\Gamma(a)\Gamma(b)} \frac{t^r}{r!}. \qquad \text{[By Taylor's series of m.g.f.]}$$

This result is of little use.

19-32. *Formula for Ratio of Two Independent Gamma Variates*

If $X \sim$ gam (a, λ) and $Y \sim$ gam (b, λ) are independent variates, then (X/Y) is $B_{II}(a, b)$.

Proof. Let $u = x/y$, $v = x$, so that $x = v$, $y = v/u$. Also

$$\frac{\partial(x, y)}{\partial(u, v)} = \begin{vmatrix} \partial x/\partial u & \partial x/\partial v \\ \partial y/\partial u & \partial y/\partial v \end{vmatrix} = \begin{vmatrix} 0 & 1 \\ -v/u^2 & 1/u \end{vmatrix} = \frac{v}{u^2}.$$

Thus, $dxdy = (v/u^2)\, dudv$. The region $0 < x < \infty$, $0 < y < \infty$ in the x-y plane trasnforms to the region $0 < u < \infty$, $0 < v < \infty$ in the u-v plane. (Draw Figure).

Since X and Y are independent, their joint probability element is

$$dP(x, y) = \frac{e^{-\lambda x} x^{a-1}\, dx}{\Gamma(a) / \lambda^a} . \frac{e^{-\lambda y} y^{b-1}\, dy}{\Gamma(b) / \lambda^b} = K\lambda^{a+b} e^{-\lambda(x+y)} x^{a-1} y^{b-1}\, dx\, dy, [K^{-1} = \Gamma(a)\Gamma(b)]$$

i.e. $$dP_1 (u, v) = K\lambda^{a+b} e^{-\lambda[(1+u)/u]v} v^{a+b-1} u^{-b-1}\, du\, dv$$

By integrating **out** v we get

$$dF(u) = \frac{K\lambda^{a+b}\, du}{u^{b+1}} \int_0^{\infty} e^{-\lambda(1+u)u^{-1}v} v^{a+b-1}\, dv = \frac{K\lambda^{a+b}\, du}{u^{b+1}} . \frac{\Gamma(a+b)\, u^{a+b}}{\lambda^{a+b}.(1+u)^{a+b}}$$

i.e. $f(u)\, du = \dfrac{u^{a-1} du}{B(a, b)\, (1+u)^{a+b}}$, $0 \le u < \infty$. This shows that $U \sim B_{II}(a, b)$.

19-33. *A Special Worked-out Problem*

Example : Let $X \sim$ gam (a, λ) and $Y \sim$ gam (b, λ) be independent variates. Let

$$Z = X + Y, \quad U = \frac{X}{Y}, \quad V = \frac{X}{X+Y}, \quad W = \frac{X-Y}{X+Y}.$$

Show that Z is independent of each of U, V, W and find their p.d.f's. Deduce that

$$\mathrm{E}\left(\frac{X}{X+Y}\right) = \frac{\mathrm{E}(X)}{\mathrm{E}(X) + \mathrm{E}(Y)}; \quad \mathrm{E}\left(\frac{X-Y}{X+Y}\right) = \frac{\mathrm{E}(X) - \mathrm{E}(Y)}{\mathrm{E}(X) + \mathrm{E}(Y)}.$$

Solution. The joint distribution of X, Y is

$$dF(x, y) = \frac{\lambda e^{-\lambda x} (\lambda x)^{a-1}\, dx}{\Gamma(a)} . \frac{\lambda e^{-\lambda y} (\lambda y)^{b-1}\, dy}{\Gamma(b)}, \quad 0 < x, y < \infty \qquad ...(1)$$

Let $x = r \cos^2 \theta$, $y = r \sin^2 \theta$, so that $0 \le r < \infty$, $0 \le \theta \le \pi/2$; also

$$\frac{\partial(x,y)}{\partial(r,\theta)} = \begin{vmatrix} \cos^2\theta & \sin^2\theta \\ -2r\sin\theta\cos\theta & 2r\sin\theta\cos\theta \end{vmatrix} = 2r\sin\theta\cos\theta.$$

The joint distribution of the new variates R and θ is thus

$$dF(r,\theta) = \left(\frac{\lambda e^{-\lambda r}(\lambda r)^{a+b-1}\,dr}{\Gamma(a+b)}\right)\left(\frac{2\cos^{2a-1}\theta\sin^{2b-1}\theta\,d\theta}{B(a,b)}\right) \quad \text{....(2)}$$

It follows that $R(=Z)$ and θ are independent, and since U, V, W are function of θ only, we conclude that Z is independent each of U, V, W. Obviously $Z = R \sim \text{gam}\ (a+b, \lambda)$; $\theta \sim B_{\mathrm{I}}(a, b)$.

$$dF(\theta) = 2\cos^{2a-1}\theta\sin^{2b-1}\theta\,d\theta\,/\,B(a,b) \quad \text{...(3)}$$

(*a*) Set : $\quad u = x/y = \cot^2\theta,\ du = -2\cot\theta\,\text{cosec}^2\,\theta\ d\theta = -2(1+u)\sqrt{u}\,d\theta$

In terms of u, (3) reduces to

$$dF(u) = \frac{2}{B(a,b)}\left(\frac{u}{1+u}\right)^{\frac{2a-1}{2}}\left(\frac{1}{\sqrt{1+u}}\right)^{2b-1}\frac{du}{2(1+u)\sqrt{u}} = \frac{u^{a-1}\,du}{B(a,b)\,(1+u)^{a+b}},\ 0<u<\infty$$

Thus $U \sim B_{\mathrm{III}}\ (a, b)$.

(*b*) Set : $v = x/(x+y) = \cos^2\theta\,;\ dv = -2\sin\theta\cos\theta\,d\theta = -2\sqrt{v}\,\sqrt{1-v}\,d\theta$.

In terms of v, (3) reduces to

$$dF(v) = \frac{2}{B(a,b)}\cdot\frac{(\sqrt{v})^{2a-1}(\sqrt{1-v})^{2b-1}\,dv}{2\sqrt{v}\,\sqrt{1-v}} = \frac{v^{a-1}(1-v)^{b-1}}{B(a,b)},\quad 0<v<1.$$

Thuu $V \sim B_{\mathrm{I}}\ (a, b)$.

(*c*) Set : $w = (x-y)/(x+y) = \cos 2\theta, \quad dw = -2\sin 2\theta\ d\theta, \quad [1 + \cos 2\theta = 2\cos^2\theta, \text{ etc.}]$

In terms of w, (3) reduces to

$$dF(w) = \frac{2}{B(a,b)}(\cos^2\theta)^{(2a-1)/2}(\sin^2\theta)^{(2b-1)/2}\,d\theta = \frac{2}{B(a,b)}\left(\frac{1+w}{2}\right)^{\frac{2a-1}{2}}\left(\frac{1-w}{2}\right)^{\frac{2a-1}{2}}\frac{dw}{2\sqrt{1-w^2}}$$

$$\therefore \quad f_W(w) = \frac{1}{2B(a,b)}\left(\frac{1+w}{2}\right)^{a-1}\left(\frac{1-w}{2}\right)^{b-1},\quad -1\le w\le 1.$$

Now

$$\mathrm{E}(ZV) = \mathrm{E}(Z)\mathrm{E}(V) \Rightarrow \mathrm{E}(X) = \mathrm{E}(X+Y)\mathrm{E}[X/(X+Y)]$$

$$\mathrm{E}(ZW) = \mathrm{E}(Z)\mathrm{E}(W) \Rightarrow \mathrm{E}(X-Y) = \mathrm{E}(X+Y)\mathrm{E}[(X-Y)/(X+Y)].$$

These are equivalent to the results stated.

Problems with Solutions Provided at the End of the Text

1*. Show that, to each B_{I}-variate corresponds a *pair* of B_{III}-variates and conversely.

2*. If $X \sim \text{gam}\ (a, \lambda)$ and $Y \sim \text{gam}\ (b, \lambda)$ are independent, then the variates $U = X + Y$, $V = X/Y$ are independent, $U \sim \text{gam}\ (a+b, \lambda)$ and $V \sim B_{\mathrm{III}}\ (a, b)$.

3*. Random variables X and Y have joint p.d.f.

$$f(x, y) = 24xy, \; x>0, \; y>0, \; x+y \le 1; \; f(x,y) = 0, \text{ elsewhere}$$

Show that $U = X + Y$ and $V = X/Y$ are independent B_I and B_{II} variates.

Exercise 19(b)

1. Let X be a B_2 variate with $E(X) = 1/4$ and Var $(X) = 1/8$. Find the values of the parameters of X. What it is modal value ?
2. Let X be $B_2(a, b)$ and put $c = a + b$. (i) Show that if $a > 2$, $b > 0$, then

$$E(X^{-1}) = \frac{c-1}{a-1}, \quad \text{Var}(X^{-1}) \frac{b(c-1)}{(a-1)^2 (a-2)},$$

   (ii) Find the density of X^{-1} and $(1 - X)/X$.

   (iii) Express $E[X/(1 - X)]$ in terms of c and $\mu = E(X)$. What restrictions are required (on μ and c) in order for the formula to be valid ?

   (iv) Can $E(1/X)$ be unity ?
3. The variate X has the p.d.f. : $f(x) = ax^3/(1 + 2x^6)$, $x \ge 0$; $f(x) = 0$, otherwise. Find a and show that $Y = 2X/(1 + 2X)$ is a B_I-variate.
4. Let $f(x) = a^a b^b x^{a-1} (b + ax)^{-a-b}/B(a, b)$, $x > 0$; $f(x) = 0$ elsewhere. Show that

$$E(X^r) = \left(\frac{b}{a}\right)^r \frac{\Gamma(a+r)\Gamma(b-r)}{\Gamma(a)\Gamma(b)}; \quad \text{Var}(X) = \frac{b^2(a+b-1)}{a(b-1)^2(b-2)}; \quad \text{Mode} = \frac{b(a-1)}{a(b+1)}.$$

   Hint. $f(x) = \lambda^a x^{a-1}(1 + \lambda x)^{-a-b}/B(a, b)$, $\lambda = a/b$. To find $E(X^r)$, use $\lambda x = z$.
5. Show that the mean value of the positive square root of $B_I(a, b)$ variate is

$$\Gamma(a+\tfrac{1}{2}).\Gamma(a+b)/\Gamma(a)\Gamma(a+b+\tfrac{1}{2}).$$

6. Let X be $B_I(a, b)$. Suppose $I_x(a, b) = P(X \le x)$ is the c.d.f. of X, and write

$$B_x(a, b) = \int_0^x t^{a-1}(1-t)^{b-1}\, dt.$$

   Now establish the following :

   (i) $I_x(a, b) = B_x(a, b)/B(a, b)$. (ii) $I_x(a, b) = 1 - I_{(1-x)}(b, a)$.

   (iii) $I_x(a, 1) = x^a$, $I_x(1, b) = 1 - (1 - x)^b$.

   (iv) $(a + b) I_x(a, b) = a I_x(a + 1, b) + b I_x(a, b + 1)$.

   (v) $x I_x(a, b) - I_x(a + 1, b) + (1 - x) I_x(a + 1, b - 1) = 0$.
7. (a) Suppose $X \sim$ gam (λ, a) and $Y \sim$ gam (λ, b) are indep. Show that $U = X + Y$, $V = (X - Y)/(X + Y)$ are independent variates.

   (b) Let $f(x, y) = e^{-(x+y)} x^3 y^4/\Gamma(4)\Gamma(5)$, $x > 0$, $y > 0$. Find the p.d.f. of $U = X/(X + Y)$. What is Var (U)?
8. If $X \sim B_I(a, b)$ and $Y \sim B_I(a + \frac{1}{2}, b)$ are independent, show that $\sqrt{XY}$ is $B_I(2a, 2b)$.
9. If $X \sim$ gam (a) and $Y \sim$ gam $B_I(b, a - b)$ are independent variates, find the p.d.f. of $W = XY$.

10. Let $X_1, X_2, \ldots X_n$ be independent $B_I(a_i, b_i)$ variates. Suppose $a_i = a_{i+1} + b_{i+1}, i = 1, 2, 3, \ldots, k-1$. Show that $X_1, X_2, \ldots, X_n \sim B_I(a_k, b_1 + b_2 + \ldots + b_n)$.

11. Let $X \sim B_I(a, b)$ and $Y \sim B_{II}(p, q)$ be independent variates. Show that if $XY \sim B_I(r, s)$, then we must have $s = b + q$ and $r = a$ or $r = p$.

12. Let $X_1, X_2, \ldots, X_n$ be i.i.d. $B_I(a, 1)$ variates. Show that max $(X_1, X_2, \ldots, X_n) \sim B_I(na, 1)$.

13. If (X, Y) has density $f(x, y) = x^{a-1}(y-x)^{b-1}e^{-y}/\Gamma(a)\Gamma(b)$, $0 < x < y < \infty$, show that $W = X/Y$ has B_I distribution.

14. Show that if $X \sim B_I(n, m)$ and G is the geometric mean, then

$$\ln G = \frac{1}{B(m,n)} \frac{\partial}{\partial n} B(m,n) = \frac{\partial}{\partial n}\{\ln \Gamma(n) - \ln \Gamma(m+n)\}.$$

LAPLACE (OR DOUBLE EXPONENTIAL) DISTRIBUTION

19-40. Definition

A continuous variates X is said to possess double exponential (or Laplace) distribution, if its p.d.f. is given by

$$f(x) = \tfrac{1}{2}\lambda e^{-\lambda|x-\mu|}, \; -\infty < x < \infty, \; \lambda > 0, \; \mu < \infty. \quad \ldots(1)$$

Its special case $\mu = 0$ is most frequented ; a variate X with

$$f(x) = \tfrac{1}{2}\lambda e^{-\lambda|x|}, \lambda > 0, \; -\infty < x < \infty \quad \ldots(2)$$

is denoted by Lap (λ) : If X follows law (1), it may be denoted by Lap (μ, λ).

19-41. M.G.F., Cumulants and Absolute Moments

$$M(t:X) = e^{\mu t}\lambda^2/(\lambda^2 - t^2); \qquad M(it:X) = e^{i\mu t}\lambda^2/(\lambda^2 + t^2) \qquad [\S\ 8\text{-}16\ (15)]$$

$$K(t:X) = \ln M(t:X) = \mu t - \ln[1 - (t^2/\lambda^2)]$$

$$= \mu t + \sum_{r=1}^{\infty} \frac{(t^2/\lambda^2)^r}{r} = \mu t + \sum_{r=1}^{\infty}\left(\frac{(2r)!}{\lambda^{2r} r}\right)\frac{t^{2r}}{(2r)!}.$$

It follows that $$k_1 = \mu, \; k_{2r} = \frac{(2r)!}{\lambda^{2r}}\frac{1}{r}; \; k_{2r+1} = 0 \quad \ldots(1)$$

Thus $$k_2 = \mu_2 = 2/\lambda^2, \quad k_4 = 12/\lambda^4, \; \mu_3 = k_3 = 0, \; \mu_4 = k_4 + 3k_2^2 = 24/\lambda^4$$

$\therefore$ $$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = 0, \; \beta_2 = \frac{\mu_4}{\mu_2^2} = 6, \; \gamma_1 = 0, \; \gamma_2 = \beta_1 - 3 = 3.$$

Since $\gamma_1 = 0$, the distribution is symmetric (not skewed) and $\gamma_2 > 0$ implies that distribution is Leptocurtic. Observe that the point (β_1, β_2) lies on the straight line $\beta_2 - \beta_1 = 6$.

$$\nu_r = E(|X-\mu|^r) = \int_{-\infty}^{\infty} \frac{1}{2}\lambda e^{-\lambda|x-\mu|}.(|x-\mu|)^r\,dx = \frac{1}{2}\lambda\int_{-\infty}^{\infty} e^{-\lambda|y|}(|y|)^r\,dy, (y = |x-\mu|)$$

$$\nu_r = \lambda\int_0^{\infty} e^{-\lambda y}y^r dy = \frac{\lambda\Gamma(r+1)}{(\lambda)^{r+1}} = \frac{r!}{\lambda^r}. \qquad \text{[Even-integrand property]}$$

19-42. *Inter-quartile Range*

The Lap (μ, λ) distribution is symmetric about $X = \mu$; hence μ must be the median. Let Q_1 and Q_3 be the lower and upper quartiles; then by definition

$$\frac{1}{4} = \int_{Q_1}^{\mu} \frac{1}{2} \lambda e^{-\lambda|x-\mu|}\, dx = -\frac{1}{2}\lambda \int_0^{Q_1-\mu} e^{-\lambda|y|}\, dy, \quad [y = x-\mu]$$

$$-\frac{1}{2} = \lambda \int_0^{Q_1-\mu} e^{\lambda y}\, dy = [e^{\lambda(Q_1-\mu)} - 1], \quad [\because\ Q_1 - \mu < 0]$$

$$e^{\lambda(Q_1-\mu)} = \tfrac{1}{2} \Rightarrow \lambda(\mu - Q_1) = \ln 2 \qquad \text{...(i)}$$

Similarly, $\lambda(Q_3 - \mu) = \ln 2.$ ...(ii)

Adding (i) and (ii) we get $\lambda(Q_3 - Q_1) = 2\ln 2 \Rightarrow Q_3 - Q_1 = (2/\lambda)\ln 2.$

19-43. *A Worked-out Problem*

Example : If X_k, $1 \le k \le 4$ are i.i.d $N(0, 1)$ r.v.s., show that $Y = X_1 X_2 - X_3 X_4$ has p.d.f.

$$f(y) = \tfrac{1}{2} e^{-|y|}, \quad -\infty < y < \infty.$$

Solution. Recall that if $X_k \sim N(0, 1)$ and $X \sim$ gam (λ, a) then

$$\phi(t : X_k) = e^{-\frac{1}{2}t^2}; \quad M(t : X) = [1 - (t/\lambda)]^{-a} \qquad \text{...(1)}$$

$$\therefore \quad \phi(t : Y) = \phi(t : X_1 X_2 - X_3 X_4) = \phi(t : X_1 X_2) . \phi(-t : X_3 X_4) \qquad \text{[by indep.]}$$

$$\phi(t : X_1 X_2) = \mathrm{E}(e^{itX_1X_2}) = \mathrm{E}\,\mathrm{E}\{e^{itX_1X_2} \mid X_2 = x_2\} = \mathrm{E}\{e^{-\frac{1}{2}t^2x_2^2}\} \qquad \text{[Double-E Rule]}$$

$$= \mathrm{E}[e^{-t^2X}] = (1+t^2)^{-1/2} \qquad [\because\ X = \tfrac{1}{2}X_2^2 \text{ is gam } (\tfrac{1}{2}, 1) \text{ Exp. 19-3}]$$

Thus $\phi(t : Y) = (1+t^2)^{-1/2} . (1+t^2)^{-1/2} = 1/(1+t^2) \Rightarrow Y \sim \text{Lap } (1, 0).$

Problems with Solutions Provided at the End of the Text

1*. If X is Lap $(0, \lambda)$ and $Y = a + bX + cX^2$, find Corr (X, Y).

2*. For a random variable X with p.d.f.

$f(x, \theta) = [2\Gamma(1+\theta)]^{-1} \lambda^{\theta} \exp(-\lambda|x|^{1/\theta})$, $\theta > 0$, $\lambda > 0$, $-\infty < x < \infty$, obtain μ_r (central moments) and hence show that $\beta_2 = \Gamma(5\theta)\Gamma(\theta)/[\Gamma(3\theta)]^2$.

Exercise 19(c)

1. A r.v. has Lap (μ, λ) distribution. If c is any pos. number less than m, show that $\mathrm{E}(|X - c|) = \mu^{-1}[e^{-\lambda(\mu - c)} + \lambda(\mu - c)]$.

2. Let X and Y be two i.i.d. variates with p.d.f. $f(x) = \frac{1}{2}e^{-|x|}$, $-\infty < x < \infty$.

   (a) Find the distribution of $X + Y$.

   (b) Find the joint characteristic function of $X + Y$ and $X - Y$.

   Deduce : Var $(X + Y \mid X - Y) = 2 + (x-y)^3 (x - y + 3)/3(x - y + 1)$.

3. Let X and Y be i.i.d. variates with p.d.f. $f(x) = \frac{1}{2} \exp(-|x-\mu|)$, $-\infty < x < \infty$.
Find the Ch. Function of $Z = (1-p)X + pY$, $0 < p < 1/2$.
4. Calculate the semi-interquartile range for Lap (μ, λ) distribution

CAUCHY DISTRIBUTION

19-50. *Definition*

A random variable X having the p.d.f.

$$f(x) = b/\pi[b^2 + (x-a)^2], \quad -\infty < x < \infty, \quad b > 0. \qquad \ldots(1)$$

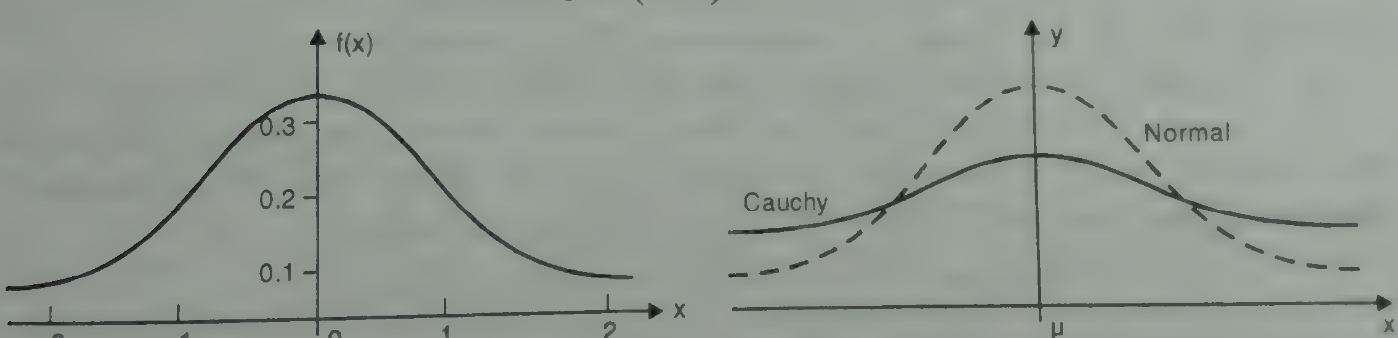
is said to be Cauchy distributed with *location* parameter a and *spread* parameter $b > 0$. A variate with density (1) is written $X \sim$ Chy (a, b). The case $a = 0$, $b = 1$, occurs frequently

$$f(x) = 1/\pi(1+x^2), \quad -\infty < x \infty \qquad \ldots(2)$$

and we write $X \sim$ Chy (0, 1) for this standard density.

Cumulative Distribution Function

$$F(x) = \int_{-\infty}^{x} \frac{b}{\pi} \frac{dt}{b^2 + (t-a)^2} = \frac{1}{2} + \frac{1}{\pi} \tan^{-1}\left[\frac{x-a}{b}\right].$$



CAUCHY DISTRIBUTION

19-51. *Slow Convergence to Zero*

The general shape of Cauchy distribution is similar to Normal curve, but it decreases more slowly for large values of x. To be specific, using c.d.f. we observe that

(i) $P\{|X-a| < b\} = F(a+b) - F(a-b) = (\frac{1}{2} + \pi^{-1} \tan^{-1} 1) - [\frac{1}{2} + \pi^{-1} \tan(-1)] = \frac{1}{2}$.

(ii) $P\{|X-a| < 2b\} = F(a+2b) - F(a-2b) = 2\pi^{-1} \tan^{-1}(2) = 0.706$.

(iii) $P\{|X-a| < 3b\} = F(a+3b) - F(a-3b) = 2\pi^{-1} \tan^{-1}(3) = 0.795$.

For $N(\mu, \sigma^2)$, the corresponding values, on the analogy in some sense that $b \sim \sigma$, are

$P\{|X-a| < \sigma\} = 0.6826$, $P\{|X-a| < 2\sigma\} = 0.9544$, $P\{|X-a| < 3\sigma\} = 0.9974$.

The higher values for $N(\mu, \sigma^2)$ indicate a larger concentration of area close to μ than for Cauchy a with implications as to their rate of slow fall off. Figure indicates this situation.

19-52. *Non-existence of Moment Generating Function*

The m.g.f. of Cauchy distribution does not exist for $t \neq 0$. To see this, consider (take $a = 0$)

$$M(t:X) = \frac{b}{\pi}\int_{-\infty}^{\infty} \frac{e^{tx}}{b^2 + x^2}\,dx \geq \frac{b}{\pi}\int_{0}^{\infty} \frac{e^{tx}}{b^2 + x^2}\,dx$$

Since $e^{tx} + 1 + tx + O(t)$, where $O(t) > 0$ for $t > 0$, the above gives

$$M(t:X) \geq \int_0^{\infty} \frac{b(1+tx)}{\pi(b^2+x^2)}\,dx \to \infty.$$

It follows that no m.g.f. for Cauchy variate exists.

19-53. *Mode and Points of Inflexion. Median*

Let $\qquad y = b/\pi(b^2 + z^2),\ z = x - a,\ -\infty < x < \infty.$

Take logarithm of both sides and then differentiate twice

$$y'/y = -2z/(b^2+z^2),\quad (y''/y) - (y'/y)^2 = -2(b^2 - z^2)/(b^2+z^2)^2,\quad (y' = dy/dz). \qquad ...(1)$$

Now $y' = 0 \Rightarrow z = 0$ and $y'' < 0$ at $z = 0$. Thus $z = 0$, i.e. $x = a$ gives the *Modal Value*. i.e. Chy (a, b) has a maximum at $x = a$.

Now put $y'' = 0$ and eliminate y' between (1i) and (1ii); this gives $3a^2 = b^2 \Rightarrow x = a \pm (b/\sqrt{3}) = a \pm 0.577b$. These are the two points of inflection of Chy (a, b) and may be compared with the two points of inflection of $N(\mu, \sigma^2)$ which are $\mu \pm \sigma$.

Note. The max. of $f(x)$ corresponds to the min. of $[f(x)]^{-1} = g(x)$ say. Now $g(x) = k[b^2 + (x-a)^2]$. $g'(x) = 2k(x-a)$, $g''(x) = 2k$. Thus $g'(x) = 0 \Rightarrow x = a$ and $g''(a) > 0$. So $g(x)$ is min. at $x = a$, i.e. $f(x)$ is max. at $x = a$. Hence $x = a$, is the Modal value of Cauchy-variate.

Median. Using the c.d.f., if m is the median of Chy (a, b), then

$$\tfrac{1}{2}F(m) = \tfrac{1}{2} + \pi^{-1}\tan^{-1}[(m-a)/b] \Rightarrow m = a.$$

19-54. *Characteristic Function :* $\phi(t:X) = e^{it\mu a - b|t|}$ $\qquad$ [§9-70]

Reproductive Property. If $X_1 \sim$ Chy (a_1, b_1) and $X_2 \sim$ Chy (a_2, b_2) are independent, then

$$(X_1 + X_2) \sim \text{Chy}\ (a_1 + a_2,\ b_1 + b_2). \qquad [\S 9\text{-}71]$$

The result for $S_n = X_1 + X_2 + ... + X_n$, follows by Mathematical Induction.

Distribution of the Mean. If $X_1, X_2, ..., X_n$ are i.i.d. Chy (a, b)-variates, then

$$\overline{X} = (X_1 + X_2 + ... + X_n)/n \sim \text{Chy}\ (a, b). \qquad [\S 9\text{-}71\ \text{Cor}]$$

Remark. There is no convergence of the arithmetic mean to a constant, not because the mean $E(X)$ does not exist, but that $\phi(t:\overline{X})$ implies it.

19-55. *Some Important Theorems Concerning Cauchy Distribution*

Theorem 1. Random variable X is Chy (a, b) iff X^{-1} is Chy $[a/(a^2 + b^2),\ b/(a^2 + b^2)]$.

Proof. Since X is Chy (a, b), its elemental probability differential is

$$dP_1(x) = b\,dx/\pi[b^2 + (x-a)^2],\ -\infty < x < \infty. \qquad ...(1)$$

Now let $Y = X^{-1}$, and put $y = 1/x$, $dy = -dx/x^2$. Then (1) reduces to

$dP_2(y) = b\,dy/\pi[b^2y^2 + (1-ay)^2]$, $-\infty < y < \infty$

$$= \frac{b}{\pi(a^2+b^2)}\frac{dy}{[y^2 - 2ay/(a^2+b^2) + 1/(a^2+b^2)]} = \frac{b'}{\pi}\frac{dy}{[(y-a')^2 + b'^2]}, \qquad \ldots(2)$$

where $b' = b/(a^2 + b^2)$, $a' = a/(a^2 + b^2)$ and $-\infty < y < \infty$.

From (2) follows the result $Y \sim$ Chy (a', b'). Now let $Y \sim (a', b')$; then $X = Y^{-1} \sim$ Chy (a'', b''), where $a'' = a'/(a'^2 + b'^2) = a$; $b'' = b'/(a'^2 + b'^2) = b$.

Thus $X \sim$ Chy $(a\ ;\ b) \Leftrightarrow X^{-1} \sim$ Chy $[a/(a^2 + b^2), b/(a^2 + b^2)]$.

Remark. This theorem *does not characterize* the Cauchy distribution.

That is, if X and X^{-1} have the same p.d.f. ; X need not be Chy (0, 1),

Let $\qquad f(x) = 1/(1 + x)^2,\ x \geq 0,\ f(x) = 0$, otherwise.

Then $Y = X^{-1}$, has p.d.f. $g(y) = 1/(1 + y)^2$, $y \geq 0$, $g(y) = 0$, otherwise.

Theorem 2. If $X \sim N(0, \sigma^2)$ and $Y \sim N(0, \sigma^2)$ are independent variates, then $U = X/Y$ is standard Cauchy distributed :

$$f(u) = 1/\pi(1 + u^2), \quad -\infty < u < \infty.$$

The converse to this theorem is not true.

Proof. Since X and Y are indep. $N(0, \sigma^2)$ variates, their joint elemental p.d.f. is

$$dP(x, y) = f_1(x)\,dx\,.\,f_2(y)\,dy = (2\pi\sigma^2)^{-1}\exp[-(x^2+y^2)/2\sigma^2]\,dx\,dy, \quad -\infty < x, y < \infty.$$

Let $u = x/y$, $v = y$ so that $x = uv$, $y = v$.

$$\frac{\partial(x, y)}{\partial(u, v)} = \begin{vmatrix} \partial x/\partial u & \partial x/\partial v \\ \partial y/\partial u & \partial y/\partial v \end{vmatrix} = \begin{vmatrix} v & u \\ 0 & 1 \end{vmatrix} = v.$$

So $dx\,dy = |\,v\,|\,du\,dv$. The elemental p.d.f. of U and V is thus

$$dP(u, v) = (2\pi\sigma^2)^{-1}\exp[-(1+u^2)v^2/2\sigma^2]\,|\,v\,|\,du\,dv, \quad -\infty < u, v < \infty$$

The marginal p.d.f. of U, obtained by summing **out** v is

$$dF(u) = \frac{du}{2\pi\sigma^2}\int_{-\infty}^{\infty}\exp[-(1+u^2)v^2/2\sigma^2]\,|\,v\,|\,dv = \frac{du}{\pi\sigma^2}\int_0^{\infty}\exp[-(1+u^2)v^2/2\sigma^2]\,v\,dv$$

$$= \frac{du}{\pi(1+u^2)}\int_0^{\infty} e^{-w}\,dw = \frac{du}{\pi(1+u^2)}, \qquad \left[w = \left(\frac{1+u^2}{2\sigma^2}\right)v^2\right]$$

Thus $\qquad f(u) = 1/\pi(1+u^2), \quad -\infty < u < \infty.$

Converse. We establish : If X and Y are i.i.d. variates then (X/Y) is standard Cauchy, but X and Y need not be normally distributed.

Proof. Let $f_1(x) = (\sqrt{2}/\pi)(1+x^4)^{-1}$, $|\,x\,| < \infty$, then the joint elemental p.d.f. of i.i.d. variates is

$$dP(x, y) = \frac{2}{\pi^2} \cdot \frac{dx\,dy}{(1+x^4)(1+y^4)}, \quad -\infty < x, y < \infty.$$

Under the transformations, $u = x/y$, $v = y$; $dx\,dy = |v|\,du\,dv$, the elemental p.d.f. becomes

$$dP(u, v) = \frac{2}{\pi^2} \frac{|v|\,du\,dv}{(1+v^4)(1+u^4v^4)}, \quad -\infty < u, v < \infty.$$

The marginal p.d.f. of U is thus

$$dF(u) = \frac{2\,du}{\pi^2} \int_{-\infty}^{\infty} \frac{|v|\,dv}{(1+v^4)(1+u^4v^4)} = \frac{4\,du}{\pi^2 u^4} \int_0^{\infty} \frac{v\,dv}{(1+v^4)(v^4+w^2)}, \qquad \left[w = \frac{1}{u^2}\right]$$

$$= \frac{2w^2\,du}{\pi^2} \int_0^{\infty} \frac{dz}{(1+z^2)(w^2+z^2)} = \frac{2w^2\,du}{\pi^2(1-w^2)} \int_0^{\infty} \left(\frac{1}{z^2+w^2} - \frac{1}{1+z^2}\right) dz, \quad [v^2 = z]$$

$$= \frac{2w\,du}{\pi^2(1-w^2)} \left\{\frac{1}{w}\tan^{-1}\left(\frac{z}{w}\right) - \tan^{-1} z\right\}_0^{\infty} = \frac{w\,du}{\pi(1+w)} = \frac{du}{\pi(1+u^2)}.$$

Thus, $U = X/Y$ is Chy (0, 1), but X and Y are not $N(0, 1)$ variates.

Note. §20-72 contains proof for Correlated r.v.'s.

Cor. Folded Cauchy Distribution :

If $Z = X/Y$, then $|Z| = |X/Y|$ is called *folded* Cauchy variate. Now p.d.f. of Z is

$$f(z) = 1/\pi(1+z^2), \quad -\infty < z < \infty$$

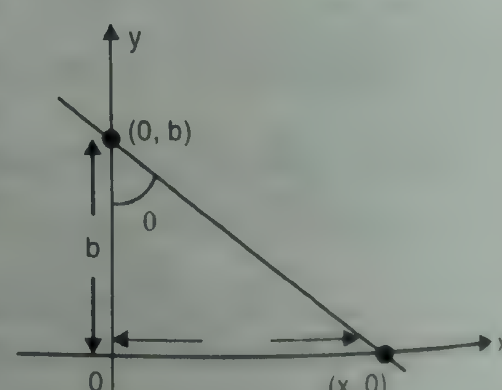
If $g(z)$ is the p.d.f. of $|Z|$, then $g(z) = f(z) + f(-z)$, $z > 0$; hence

$$g(z) = 2/\pi(1+z^2), \quad z > 0. \quad \text{[Folded-Cauchy density]}$$

Theorem 3. If $X \sim N(0, \sigma_1^2)$, and $Y \sim N(0, \sigma_2^2)$ are independent variates, then (X/Y) and $X/|Y|$ have the same standard Cauchy distribution [Example 16-27].

19-56. *Worked-out Problems*

Example 1. A needle spins about the point $(0, b)$ of the x-y plane, with $b > 0$ and comes to stop thereby determining an angle θ. The direction of the needle then intersects OX at a point $(X, 0)$. Assuming θ is uniformly distributed over $(-\pi/2, \pi/2)$, find the p.d.f. of the r.v. X.



Solution. We are given $f(\theta) = 1/\pi$, $|\theta| < \pi/2$; $f(\theta) = 0$, $|\theta| > \pi/2$.

The r.v. X is given by $X = b$ to $\tan\theta$. Thus

$$F_X(x) = P\{X \le x\} = P\{b\tan\theta \le x\} = P\{\theta \le \tan^{-1}(x/b)\} \quad [\theta_1 = \tan^{-1}(x/b)]$$

$$= \int_{-\pi/2}^{\theta_1} \frac{d\theta}{\pi} = \frac{1}{\pi}\left(\theta_1 + \frac{\pi}{2}\right) = \frac{1}{2} + \frac{1}{\pi}\tan^{-1}\left(\frac{x}{b}\right).$$

Differentiating : $\quad f_X(x) = \dfrac{1}{\pi} \cdot \dfrac{b}{b^2 + x^2}. \qquad$ [i.e. $X \sim$ Chy $(0, b)$]

***Example* 2.** If X denotes the tangent of an angle θ (measured in radians) chosen at random from $(-\pi/2, \pi/2)$, then X is Cauchy distributed.

Solution. Here $dP_1(\theta) = d\theta/\pi, -\pi/2 < \theta < \pi/2$. Now $X = \tan \Theta$, so put $x = \tan \theta$, $dx = \sec^2 \theta \, d\theta = (1 + x^2) \, d\theta$, so that $d\theta = dx/(1 + x^2)$.

Further, $-\pi/2 < \theta < \pi/2 \Rightarrow -\infty < x < \infty$; hence the probability differential for X is

$$dP(x) = dx/\pi(1 + x^2), -\infty < x < \pi.$$

Remark. The same result shall be obtained if θ is $U(0, \pi)$ or θ is $U(0, 2\pi)$.

***Example* 3.** The random variables X and Y are indep. X is Pois (m) and Y is Chy$(0, a)$. Show that the characteristic function of XY is $\exp \{m(e^{-a|t|} - 1)\}$.

Solution. Recall : $M(t : X) = e^{m(e^t - 1)}, \phi(\theta : Y) = e^{-a|\theta|}$ ...(1)

$$\therefore \quad \phi(t : XY) = \mathsf{E}(e^{itXY}) = \mathsf{E}_X[\mathsf{E}_Y(e^{itXY} \mid X = x)\} = \mathsf{E}_X\{e^{-a|tX|}\} = \mathsf{E}\{e^{-a|t|X}\}$$

$$= \exp[m(e^{-a|t|} - 1)\}, \qquad \text{[by } 1(b)]$$

Exercise 19(d)

1. Discuss : Cauchy distribution is often quoted as a distribution providing counter examples.
2. (i) If $X \sim$ Chy $(0, 1)$, find $P\{X^2 < c\}$.
   (ii) $X \sim$ Chy $(0, 1)$, find the values a and b such that
   $$P\{|X| < a\} = \tfrac{1}{2}, P\{|X| < b\} = 0.6828.$$
3. Let $X \sim$ Chy $(0, 1)$. Find the p.d.f. of (i) $Y = a + bX, b \neq 0$
   (ii) $U = 1 - X^3$, (iii) $W = 1/X$, (iv) $Z = X^2$, identify the distribution.
4. If X_1, X_2, X_3, X_4 are independent $N(0, 1)$ variates, find the distribution of
   (a) $(X_1/X_2) + (X_3/X_4)$. (b) $(X_2/|X_1|) + (X_2/|X_2|)$.
5. Let $X_1, X_2, ..., X_n$ be independent variates with $X_i \sim$ Chy (μ_i, λ_i), $i = 1, 2, ..., n$. Show that $X = [\Sigma(1/X_i)]^{-1} \sim$ Chy (μ', λ') where
   $$\mu' = \frac{\mu}{\mu^2 + \lambda^2}, \lambda' = \frac{\lambda}{\lambda^2 + \mu^2}, \lambda = \sum_{i=1}^{n}\left(\frac{\lambda_i^2}{\lambda_i^2 + \mu_i^2}\right), \mu = \sum_{i=1}^{n}\left(\frac{\mu_i^2}{\lambda_i^2 + \mu_i^2}\right).$$
6. If $X_i, 1 \le i \le n$ are i.i.d. Chy $(0, 1)$ variates, find the p.d.f. of $Y = \{\Sigma(a_iX_i + b_i)]^{-1}$ where $a_i \neq 0$ and b_i are real numbers.
7. If X be Chy $(0. 1)$ show that for every $-\infty < y < z < \infty$, $\mathsf{E}(2X \mid y < X \le z) = a/b$, where $a = \ln[(1 + x^2)/(1 + y^2)]$, $b = \tan^{-1}[(z - y)/(1 + zy)]$.
8. Let X and Y be i.i.d. Chy $(0, a)$ variates. Prove that $U = XY$ has the density $\{a^2/\pi^2(u^2 - a^4)\} \ln(u^2/a^4)$, $-\infty < u < \infty$.
   Show further that, if $a = 1$, the p.d.f. of X/Y is identical with the p.d.f. of XY and explain why it is so.
9. If X, Y, Z are i.i.d. Chy $(0, a)$ variates, find the p.d.f. of XYZ.
10. If X is Chy $(0, 1)$, find $\mathsf{E}\{\min(|X|, 1)\}$.

11. $\overline{X}_n$ is the mean of n independent variates distributed like X, and X has a symmetric distribution. If $\overline{X}_n$ has exactly the same distribution as X for all n, prove that the characteristic function of X is $f_X(t) = e^{-c|t|}$ for some real constant c. Identify X and generalize the above result.

12. Let $X_1, X_2, \ldots$ be a sequence of independent variates. If n is perfect square, then X_n is Chy (0, 1), otherwise X_n has a c.d.f. $F(x)$, with mean zero and variance σ^2. Discuss the asymptotic behaviour of $S_n/\sqrt{n}$.

19-60. *Miscellaneous Worked-out Problems*

***Example* 1.** If X is a Weibull variate, find its mean and variance.

Solution. Here $f(x) = (k/a)\,[(x-\mu)/a]^{k-1} \exp\{-[(x-\mu)/a]^k\},\ x > \mu,\ a > 0,\ k > 0.$ ...(1)

$$\therefore \quad E(X-\mu)^r = \frac{k}{a}\int_\mu^\infty \left(\frac{x-\mu}{a}\right)^{k-1} (x-\mu)^r \exp\left[-\left(\frac{x-\mu}{a}\right)^k\right] dx$$

$$= ka^r \int_0^\infty y^{k+r-1} \exp(-y^k)\, dy \qquad \left[\frac{x-\mu}{\sigma} = y\right]$$

$$= a^r \int_0^\infty e^{-z} z^{(r/k)}\, dz = a^r\, \Gamma[1+(r/k)]. \qquad [z = y^k] \qquad \ldots(2)$$

Put $r = 1, 2$, in $E(X-\mu)^r = a^r\,\Gamma[1+(r/k)]$ to obtain :

$$E(X-\mu) = a\Gamma(1+1/k) \Rightarrow E(X) = \mu + a\Gamma(1+1/k).$$

$$E(X-\mu)^2 = a^2\,\Gamma(1+2/k) \Rightarrow E(X^2) = 2\mu[\mu + a\Gamma(1+1/k)] - \mu^2 + a^2\Gamma(1+2/k).$$

$$\therefore \quad \text{Var}(X) = E(X^2) - E^2(X) = a^2\,\{\Gamma(1+2/k) - \Gamma(1+1/k)]^2\}.$$

Note. For the standard Weibull distribution ($\mu = 0$, $a = 1$), $\mu'_r = \Gamma[(r/k)+1]$
Thus $E(X) = \Gamma(1+k^{-1})$; $\text{Var}(X) = \Gamma(1+2k^{-1}) - [\Gamma(1+k^{-1})]^2$, etc.

Comments. For Weibull distribution $f(x) = \alpha\beta\, x^{\beta-1}\, e^{-\alpha x^\beta},\ x > 0$

$$\mu = \alpha^{-1/\beta}\,\Gamma\left(1+\frac{1}{\beta}\right), \quad \sigma^2 = \alpha^{-2/\beta}\left\{\Gamma\left(1+\frac{2}{\beta}\right) - \Gamma^2\left(1+\frac{1}{\beta}\right)\right\}.$$

***Example* 2.** Find the m.g.f. and moments of Gumbel (or extreme value) distribution.

Solution. Gumbel distribution : $F(x) = \exp[-e^{-\lambda(x-a)}],\ \lambda > 0,\ |a| < \infty,\ -\infty < x < \infty.$

Let $z = e^{-\lambda(x-a)}$, so that $f(x) = F'(x) = \lambda z e^{-z}$, $dx = -(dz/\lambda z)$. $-\infty < z < \infty$.

$$\therefore \quad E(X-a)^k = \int_{-\infty}^\infty (x-a)^k f(x)\, dx = \int_0^\infty \left(\frac{\ln z}{-\lambda}\right)^k . e^{-z}\, dz = (-1/\lambda)^k \int_0^\infty (\ln z)^k e^{-z}\, dz. \qquad \ldots(1)$$

Now $\Gamma(n) = \int_{-\infty}^\infty e^{-t} t^{n-1}\, dt$, $\Gamma'(n) = \int_0^\infty e^{-t} t^{n-1} \ln t\, dt$, ..., $\Gamma^{(k)}(n) = \int_0^\infty e^{-t} t^{n-1} (\ln t)^k\, dt$

Thus $$\Gamma^{(k)}(1) = \int_0^\infty e^{-t} (\ln t)^k\, dt \qquad \ldots(2)$$

$$\therefore \quad E(X-a)^k = (-1/\lambda)^k . \Gamma^{(k)}(1). \qquad \ldots(3)$$

The psi function (digamma function) is defined by $y(x) = d\Gamma(x)/dx$; and its nth derivative is

$$\psi^{(n)}(x) = n!(-1)^{n+1} \sum_{r=0}^\infty \frac{1}{(x+r)^{n+1}} \Rightarrow \Gamma^{(k)}(1) = \Psi^{(k-1)}(1) = (k-1)!(-1)^k \sum_{r=0}^\infty \frac{1}{(1+r)^k}.$$

Thus $$E(X-a)^k = (k-1)!(1/\lambda)^k \sum_{r=0}^\infty \frac{1}{(1+r)^k}. \qquad \ldots(4)$$

Recall : $-\Gamma'(1) = c = \lim_{n\to\infty}\left(1+\frac{1}{2}+\frac{1}{3}+\ldots+\frac{1}{n}-\ln n\right) = 0.57216.$ [Euler's constant]

So $\mu = E(X) = a - \Gamma'(1)/\lambda = a + (c/\lambda)$ [by (2)]

$\text{Var } X = E(X-a)^2 - (\mu - a)^2 = (1/\lambda^2)(\Sigma 1/n^2) - (c/\lambda)^2 = [(\pi^2/6) - c^2]/\lambda^2.$

$M(t:X) = E(e^{tX}) = \int_{-\infty}^{\infty} e^{tX} f(x)\,dx = e^{at}\int_0^{\infty} e^{-z} z^{-t/\lambda}\,dz = e^{at}\Gamma\left(1-\frac{t}{\lambda}\right),\quad [t<\lambda].$

Problems with Solutions Provided at the End of the Text

1*. The r.v. X has Weibull law : $f(x) = (6\sqrt{x})^{-1}\exp[-\sqrt{x}/3], x \geq 0$ and r.v. Y has the Weibull law :
$$g(y) = (4\sqrt{y})^{-1}\exp(-\sqrt{y}/2), y \geq 0.$$
Find the density of $Z = X/Y$ and hence show that $P(X < Y) = 2/5$.

2*. Let $X_1, X_2, \ldots, X_n$ be a random sample from Weibull distribution
$$f(x) = \frac{k}{a}\left(\frac{x-\mu}{a}\right)^{k-1}\exp\left[-\left(\frac{x-\mu}{a}\right)^k\right];\quad x>\mu,\ (a>0,\ k>0). \qquad \ldots(1)$$
Find the distribution of $Y = \min\{X_1, X_2, \ldots, X_n\}$ and identify it.

3*. Let $X_1, X_2, \ldots, X_n$ be i.i.d. expo (λ) variates. Let
$$Y = X_1 + \tfrac{1}{2}X_2 + \ldots + (1/n)X_n;\ Z = \max\{X_1, X_2, \ldots, X_n\}.$$
Prove that Y and Z have the same distribution.

4*. A family of distributions is defined by
$$(df/dx) = xf/(b_0 + b_2x^2 + b_4x^4) \qquad \ldots(1)$$
and the frequency function vanishes at the terminals of its range. Show that the central moments are given by
$$b_0(2n+1)\,\mu_{2n} + b_2(2n+3)\,\mu_{2n+2} + b_4(2n+5)\,\mu_{2n+4} = -\mu_{2n+2}.$$

5*. Show that for a Pearson distribution $df/f = (a+x)\,dx/(b_0 + b_1x + b_2x^2)$ the characteristic function ϕ obeys the relation
$$b_2\theta\phi'' + (1 + 2b_2 + b_1\theta)\,\phi' + (a + b_1 + b_0\theta)\phi = 0$$
where $\theta = it$, $\phi' = df/d\theta$, etc. Deduce the recurrence relation for moments. Show also that the cumulants generating function ψ obeys the relation
$$b_2\theta(\psi'' + \psi'^2) + (1 + 2b_2 + b_1\theta)\psi' + (a + b_1 + b_0\theta) = 0.$$
Hence show that the cumulants obey the recurrence relation
$$rb_2\left[\binom{r-1}{1}k_2k_{r-1} + \ldots + \binom{r-1}{j}k_{j+1}k_{r-j} + \ldots + \binom{r-1}{r-2}k_{r-1}k_2\right]rb_1k_r + [1+(r+2)b_2]k_{r+1} = 0.$$

Assorted Problems

1. Six independent observations are to be made on a r.v. with density $f(x) = 2x$, $0 < x < 1$. The interval (0, 1) is divided into four class intervals of equal length. What is the probability that exactly three observations fall in the left-most class interval, two in the next, one in the further next and none in the right-most class interval ?
2. If $X \sim U(0, 1)$ and $Y \sim f_y$ are independent, find the density of $Z = X - Y$, where $f(y) = 2y$, $0 < y < 1$. Evaluate $P(Y > X)$.
3. If $X \sim B_I(a, b)$ and $Y \sim \text{gam}\,(1, a+b)$ are independent, show that XY and $(1 - X)Y$ are independent.
4. Let $X \sim U(0, 1)$. Find a transformation $u = u(x)$ such that
   (i) $Y = u(X)$ is gam (1/2, 1/2) (ii) $Y = u(X)$ is $B_I(2, 1)$.
5. Let $X \sim B_{II}(2, 2)$. Find a transformation $u = u(x)$ such that $Y = u(X)$ has the p.d.f.
   $$f(y) = 12y^3(1 - y^2),\quad 0 < y < 1;\ f(y) = 0, \text{ elsewhere.}$$
6. The daily output of a perishable commodity is described by gam (1, 4) variate, and the proportion of this output which perishes before reaching the market is described by an independent $B_{II}(2, 2)$ variate. Show that the probability that a day's production will result in the loss of at most 3 units and atleast 2 units reaching the market is $3(e^3 - 4)/e^5$.
7. Variate X has the p.d.f. (called *arc-sine* density) $f(x) = 1/\pi\sqrt{1 - x^2}$), $|x| < 1$.

   Show that : (a) X^2 and $\frac{1}{2}(1 + X)$ are identically distributed (i.d.). (b) X^2 and $1 - X^2$ are i.d., (c) X and $2X(1 - X^2)^{1/2}$ are identically distributed.
8. Let X and Y be i.i.d. $f(x) = \{\pi\sqrt{(4/b^2) - x^2}\}^{-1}, |x| < |2/b|$. Show that $(X + Y)/b$ and XY are identically distributed.
   Conversely, let X and Y be two i.i.d. variates such that :
   (i) $(X + Y)/b$ and XY are identically distributed.
   (ii) The common c.d.f. of X and Y, is symmetric about origin and is uniquely determined by the moments. Show that X and Y follow the above law.
9. Let $f(x, y) = \dfrac{x^{a-1}\, y^{b-1}(1 - x - y)^{c-1}}{\Gamma[(a)\,\Gamma(b)\,\Gamma(c)/\Gamma(a+b+c)]}$, $x > 0, y > 0, x + y < 1$.
   [*Dirichlet's distribution* with parameters a, b, c and written Dirch (a, b, c)]
   (i) Show that if (X, Y) is Dirich (a, b, c) then X and $Y/(1 - X)$ are independent and find their respective distributions.
   (ii) Show that if (X, Y) is Dirich $(a, a + 1, a)$, then $E[(X - X^2)/Y] = 2a/(3a + 1)$.
10. If $X_i \sim \text{gam}\,(a_i)$, $i = 1, 2, 3$ are independent variates, show that $X_1/(X_1 + X_2 + Y_3)$ and $X_2/(X_1 + X_2 + X_3)$ follow a bivariate Dirichlet distribution.
11. A random vector (X, Y) has Dirich (a, b, c) distribution. Show that X is $B_I(a, b + c)$ and Y is $B_I(b, a + c)$, prove that $\rho(X, Y) < 0$ and explain why it is so.
12. If X is gam (p), Y is $B_I(p, p - q)$ then $\sqrt{2XY}$ is $N(0, 1)$ if $q = \frac{1}{2}$. However, if $q = p - q = 1$, then XY is Expo (1).
13. For the Pearsonian family of distributions specified by $f'/f = (x + a)/(b_0 + b_1x + b_2x^2)$ show that
   $$\frac{\text{Mean} - \text{Mode}}{\text{S.D.}} = \frac{\sqrt{\beta_1}\,(\beta_2 + 3)}{(5\beta_2 - 6\beta_1 - 9)}.$$

It is better to live rich than to die rich. ***(Samuel Johnson)***

Everyone has a photographic memory. Some don't have film.

Appendix : Compound Distributions

F

F-1. *Definition*

Let $F(x, \theta_1, \theta_2, \ldots, \theta_k)$ be a given distribution function. If any of its parameters is ascribed a distribution (the compounding distribution), then the unconditional distribution of $F(x, \theta_1, \theta_2, \ldots, \theta_k)$ is called a *compound distribution.*

F-2. *Compound Poisson Distribution*

A variate X is Pois (M) where the parameter M itself is gam (λ, r) with r an integer ; to find the unconditional distribution of X. Here,

$$P\{X = k \mid M = m\} = e^{-m} m^k / k!;\ m \geq 0,\ k = 0, 1, 2, \ldots;\ f(m) = \lambda^r e^{-\lambda m} m^{r-1} / \Gamma(r),\ m \geq 0$$

The Multi-Stage (Break-down) Rule for continuous variates is

$$P\{X = k) = \Sigma_i P(m_j)\, P(X = k \mid m_j) = \int_{-\infty}^{\infty} f(m)\, P(X = k \mid M = m)\, dm$$

$$\therefore \quad P(X = k) = \int_0^{\infty} \frac{e^{-m}\, m^k}{k!} \cdot \frac{\lambda^r m^{r-1}\, e^{-\lambda m}}{\Gamma(r)}\, dm = \frac{\lambda^r}{k!\,\Gamma(r)} \int_0^{\infty} m^{r+k-1} e^{-(1+\lambda)m}\, dm$$

$$= \frac{\lambda^r}{k!\,\Gamma(r)} \frac{\Gamma(r+k)}{(1+\lambda)^{r+k}} = \frac{(r+k-1)!}{k!\,(r-1)!} \left(\frac{\lambda}{1+\lambda}\right)^r \left(\frac{1}{1+\lambda}\right)^k = \binom{r+k-1}{k} p^r q^k$$

where $p = \lambda/(1 + \lambda)$. Thus, the unconditional distribution of X is Neg-bin (r, p).

F-3. *Compound Binomial Distribution*

A variate X is bin (N, p) where the parameter N itself is Pois (λ); to find the unconditional distribution of X. Here

$$P(X = x \mid n) = \binom{n}{x} q^{n-x} p^x,\ x = 0, 1, 2, \ldots, n;\ f(n) = e^{-\lambda} \lambda^n / n!\, n = 0, 1$$

The Multi-Stage (Break-down) Rule : $P(X = x) = \sum_n P(n)\, P(X = x \mid n)$ yields

$$\therefore \quad P(X = x) = \sum_{n=0}^{\infty} \frac{e^{-\lambda}\lambda^n}{n!} \frac{n!\, p^x\, q^{n-x}}{x!\,(n-x)!} = \frac{e^{-\lambda}(\lambda p)^x}{x!} \sum_{n=x}^{\infty} \frac{(\lambda q)^{n-x}}{(n-x)!}$$

$$= \frac{e^{-\lambda}(\lambda p)^x}{x!} e^{\lambda q} = \frac{e^{-\lambda p}(\lambda p)^x}{x!},\ x = 0, 1, 2, \ldots \infty.$$

Thus, the unconditional distribution of X is Pois (λp).

F-4. *Hyp-geom and Bin Combination Theorem*

If Y is bin (N, p) and ($X \mid y$) is H-G (N, y, n), $y = 0, 1, 2, \ldots, N$, then prove that X is bin (n, p) and ($Y \mid x$) is x + bin ($N - n$, p); i.e.

$$P(Y = y \mid X = x) = \binom{N-n}{y-x} p^{y-x} q^{N-n-(y-x)}.$$

Proof. $P(X = x \mid Y = y) = \binom{y}{x}\binom{N-y}{n-x}\Big/\binom{N}{n}$ (H-G-pmf)

$$P(X = x, Y = y) = \frac{\binom{y}{x}\binom{N-y}{n-x}}{\binom{N}{n}}\binom{N}{y}q^{N-y}p^{y}. \qquad [\because \; Y \sim \text{bin}\,(N, p)]$$

$$= \frac{y^{(x)}}{x!}\frac{(N-y)^{(n-x)}}{(n-x)!}\frac{N^{(y)}}{y!}\frac{n!}{N^{(n)}}q^{N-y}p^{y} = \binom{n}{x}\binom{N-n}{y-x}p^{y}q^{N-y}.$$

Since $x \le y$ and $x \ge n + y - N$, so range of y is $x \le y \le N - n + x$.

To find the marginal density of X, we sum **out** y. Putting $y = x + z$, we get

$$P(X = x) = \binom{n}{x}\sum_{y=x}^{N-n+x}\binom{N-n}{y-x}p^{y}q^{N-y} = \binom{n}{x}p^{x}q^{n-x}\sum_{z=0}^{N-n}\binom{N-n}{z}p^{z}q^{N-n-z}$$

$$= \binom{n}{x}p^{x}q^{n-x}(q+p)^{N-n} = \binom{n}{x}q^{n-x}p^{x}.$$

$$P\{Y = y \mid X = x\} = \frac{P\{X = x, Y = y\}}{P\{X = x\}} = \binom{n}{x}\binom{N-n}{y-x}p^{y}q^{N-y}\Big/\binom{n}{x}q^{n-x}p^{x}$$

$$= \binom{N-n}{y-x}q^{N-n-(y-x)}p^{y-x}.$$

Note.
$$\boxed{M^{(k)} = \frac{\text{Base !}}{(\text{Base} - \text{Power})\,!} = \frac{M\,!}{(M-k)\,!}}$$

F-5. *Normal-Normal Combination Theorem*

If $Y \sim N(X, w)$ and $X \sim N(m, v)$, show that $Y \sim N(m, v + w)$ and

$$(X \mid Y = y) \sim N\left(\frac{vy + wm}{v + w}, \frac{vw}{v + w}\right).$$

Proof.
$$M(t : Y) = \mathsf{E}(e^{tY}) = \mathsf{E}_X(\mathsf{E}_Y e^{tY} \mid X) = \mathsf{E}(e^{Xt + \frac{1}{2}t^2 w}) = e^{\frac{1}{2}t^2 w}\mathsf{E}(e^{tX})$$

$$= e^{\frac{1}{2}t^2 w}\exp(mt + \tfrac{1}{2}t^2 v) = \exp[mt + \tfrac{1}{2}t^2(v + w)].$$

Thus, Y is Normal with stated parameters. We now use Baye's Reversal Rule :

$$f(x \mid y) = \frac{f(y \mid x)\,f_1(x)}{f_2 y} = \frac{\exp[-\frac{1}{2}(y-x)^2/w]}{\sqrt{2\pi w}}\,\frac{\exp[-\frac{1}{2}(x-m)^2/v]}{\exp[-\frac{1}{2}(y-m)^2/(v+w)]}\,\frac{\sqrt{2\pi(w+v)}}{\sqrt{2\pi v}}$$

$$= \frac{1}{\sigma\sqrt{2\pi}}\exp\left\{-\frac{1}{2}\left[\frac{(y-x)^2}{w} + \frac{(x-m)^2}{v} - \frac{(y-m)^2}{w+v}\right]\right\}, \quad \left(\sigma^2 = \frac{wv}{w+v}\right)$$

$$= \frac{1}{\sigma\sqrt{2\pi}}\exp\left\{-\frac{1}{2}\left[\frac{x^2}{\sigma^2} - \frac{2x(vy + wm)}{vw} + \frac{y^2v^2 + w^2m^2}{wv(w+v)} + \frac{2my}{w+v}\cdot\frac{\sigma^2}{\sigma^2}\right]\right\}$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left[x^2 - \frac{2x(vy+wm)}{w+v} + \left(\frac{vy+wm}{w+v}\right)^2\right]\Big/\sigma^2\right\}. \quad [\text{use } wv = (w+v)\sigma^2]$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(x - \frac{vy+wm}{w+v}\right)^2 \Big/\sigma^2\right\}.$$

F-6. Worked-out Problems

***Example* 1.** Find the unconditional distribution of X in the following cases :

(a) $X \sim \text{bin}(N, p)$ and $N \sim \text{bin}(m, p')$.

(b) $X \sim \text{bin}(n, p)$ and $p \sim B_1(a, b)$

(c) $X \sim \text{Pois}(M)$ and $M \sim \text{gam}(a, \lambda)$.

(d) $X \sim \text{NB}(n, p)$ and $p \sim B_1(a, b)$

(e) $X \sim \text{NB}(n, p)$ and $N \sim \text{gem}(p')$.

(f) $X \sim \text{Expo}(Y)$ and $Y \sim \text{Unif}(0, 1)$.

(g) $X \sim N(0, 1/\theta)$ and $\theta \sim \text{gam}(\frac{1}{2}n, \frac{1}{2}n)$

(h) $X \sim \text{Bur}(\lambda; a, b)$ and $a \sim \text{gam}(c, \lambda)$

(i) $X \sim \text{gam}(\frac{1}{2}m, \theta)$ and $\theta \sim \text{gam}(\frac{1}{2}n, \lambda), (\lambda = n/m)$.

In each case, name the distribution if possible.

Solution. We can proceed as in §E-2 or use Generating functions.

(a) $G(t:X) = \mathsf{E}(t^X) = \mathsf{E}_N\{\mathsf{E}(t^X \mid N=n)\} = \mathsf{E}(q+pt)^N = \mathsf{E}(t')^N = (q'+p't')^m$

$= [q' + p'(q+pt)]^m = [1 - pp' + pp't]^m$. Thus $X \sim \text{bin}(m, pp')$.

(b) $P\{X=k\} = \int_0^1 P(X=k \mid p=z)\, f_p(z)\, dz = \int_0^1 \binom{n}{k} z^k (1-z)^{n-k} \cdot \dfrac{z^{a-1}(1-z)^{b-1} dz}{B(a,b)}$

$= \binom{n}{k}\int_0^1 \dfrac{z^{k+a-1}(1-z)^{n-k+b-1}}{B(a,b)} dz = \binom{n}{k}\dfrac{B(k+a, n-k+b)}{B(a,b)}, k = 0, 1, \ldots, n$ (Beta-bin distribution)

(c) $G(t:X) = \mathsf{E}(t^X) = \mathsf{E}_M\{\mathsf{E}(t^X \mid M=m)\} = \mathsf{E}[e^{M(t-1)}]$ (m.g.f. of M)

$$= \left(1 - \frac{t-1}{\lambda}\right)^{-a} = \left(\frac{\lambda}{1+\lambda}\right)^a \left(1 - \frac{1}{1+\lambda}t\right)^{-a} = \frac{p^a}{(1-qt)^a}.$$

Thus $X \sim \text{NB}(a, p)$, where $p = \lambda/(1+\lambda)$. [gamma mixture of Poissons]

(d) $P\{X=k\} = \int_0^1 P(X=k \mid p=z)\, f_p(z)\, dz = \int_0^1 \binom{k+n-1}{n-1} z^n (1-z)^k \dfrac{z^{a-1}(1-z)^{b-1}}{B(a,b)} dz$

$$= \binom{k+n-1}{n-1}\frac{1}{B(a,b)}\int_0^1 z^{a+n-1}(1-z)^{b+k-1} dz$$

$$= \binom{n+k-1}{n-1}\frac{B(a+n, b+k)}{B(a,b)}, k = 0, 1, 2, \ldots$$

For the special case, $X \sim \text{gem}(p)$, using $n = 1$, $k \to k-1$, we get

$$P(X=k) = \frac{B(a+1, b+k-1)}{B(a,b)} = \frac{a\Gamma(b+k-1)\,.\,\Gamma(a+b)}{\Gamma(a+b+k)\Gamma(b)}, \quad k = 1, 2, 3, \ldots$$

(e) $G(t:X) = \mathrm{E}(t^X) = \mathrm{E}_N\{\mathrm{E}(t^X \mid N)\} = \mathrm{E}\left(\dfrac{p}{1-qt}\right)^N = \mathrm{E}(t')^N$

$$= \frac{p't'}{1-q't'} = \frac{pp'}{1-qt-pq'} = \frac{pp'}{1-pq'}\left(1-\frac{qt}{1-pq}\right)^{-1}. \text{ Thus } X \sim \text{geom } [pp'/(1-pq')]$$

(f)
$$f_X(x) = \int_0^1 ye^{-yx}.1.dy = \left[y\left(\frac{e^{-xy}}{-x}\right)-(1)\left(\frac{e^{-xy}}{x^2}\right)\right]_0^1, \quad \text{(Integ. by Parts)}$$

$$= -e^{-x}\left(\frac{1}{x}+\frac{1}{x^2}\right)+\frac{1}{x^2},\ x>0$$

$$= F_X(t) = \int_0^t\left[\frac{1}{x^2}-e^{-x}\left(\frac{1}{x}+\frac{1}{x^2}\right)\right]dx = 1-\left(\frac{1-e^{-t}}{t^2}\right), t \geq 0.$$

(g) $$f(x) = \int_0^\infty \frac{\sqrt{\theta}\, e^{-\frac{1}{2}\theta x^2}}{\sqrt{2\pi}} \cdot \frac{e^{-(1/2)n\theta}\theta^{(1/2)n-1}(\frac{1}{2}n)^{n/2}}{\Gamma(\frac{1}{2}n)} d\theta$$

$$= \frac{(\frac{1}{2}n)^{n/2}}{\sqrt{2\pi}\,\Gamma(n/2)}\int_0^\infty e^{-\frac{1}{2}(n+x^2)\theta}\theta^{[(n+1)/2]-1}\,d\theta = \frac{(\frac{1}{2}n)^{n/2}}{\sqrt{2\pi}\Gamma(n/2)} \cdot \frac{\Gamma[(n+1)/2]}{[\frac{1}{2}(n+x^2)]^{(n+1)/2}}, \ -\infty < x < \infty$$

$$= \frac{1}{\sqrt{n}B(\frac{1}{2},\frac{1}{2}n)} \cdot \left(1-\frac{x^2}{n}\right)^{-(n+1)/2}, \quad -\infty < x < \infty.$$

This is the p.d.f. of Student-Fisher t statistic.

(h) $X \sim$ Bur (λ, a, b) means $f_\lambda(x) = ab\,\lambda^a x^{b-1}/(\lambda+x^b)^{a+1},\ 0 < x < \infty.$

$$f(x) = \int_0^a f(x \mid a)\, g(a)\, da = \int_0^\infty \frac{ab\,\lambda^a x^{b-1}}{(\lambda+x^b)^{a+1}} \frac{\lambda^c e^{-\lambda a} a^{c-1}}{\Gamma(c)} da$$

$$= \frac{b\lambda^c x^{b-1}}{(\lambda+x^b)\,\Gamma(c)}\int_0^\infty \frac{(\lambda e^{-\lambda})^a a^c}{(\lambda+x^b)^a} da \qquad \left(\text{Put } \frac{\lambda e^{-\lambda}}{\lambda+x^b} = e^{-k}\right)$$

$$= \frac{b\lambda^c\, x^{b-1}}{(\lambda+x^b)\,\Gamma(c)}\int_0^\infty e^{-ka}a^{(c+1)-1}\,da = \frac{b\lambda^c x^{b-1}}{(\lambda+x^b)\,\Gamma(c)} \cdot \frac{\Gamma(c+1)}{(k)^{c+1}}$$

$$= \frac{bc\lambda^c\, x^{b-1}}{(\lambda+x^b)}\,\frac{1}{[\ell n\,(\lambda+x^b)+\lambda-\ell n\,\lambda]^{c+1}}.$$

(i) $$f(x) = \int_0^\infty f(x \mid \theta)\, g(\theta)\, d\theta = \int_0^\infty \frac{\theta^{m/2}\, e^{-\theta x}\, x^{(m/2)-1}}{\Gamma(m/2)} \cdot \frac{e^{-\lambda\theta}\,\theta^{(1/2)n-1}\lambda^{(1/2)n}}{\Gamma(n/2)}\, d\theta$$

$$= \frac{(\lambda)^{n/2}\,.\,x^{(1/2)m-1}}{\Gamma(m/2)\,\Gamma(n/2)}\int_0^\infty e^{-(\lambda+x)\theta}\theta^{(m+n)/2-1}d\theta = \frac{\Gamma(\frac{1}{2}m+\frac{1}{2}n)\,(\lambda)^{n/2}}{\Gamma(\frac{1}{2}m)\,\Gamma(\frac{1}{2}n)} \cdot \frac{x^{(1/2)m-1}}{(\lambda+x)^{(m+n)/2}}, \ 0 < x < \infty$$

$$= \frac{(m/n)^{m/2}}{B(\frac{1}{2}m,\frac{1}{2}n)} \cdot \frac{x^{(1/2)m-1}}{[1+(mx/n)]^{(m+n)/2}}, 0 < x < \infty.$$

This is p.d.f. of Fisher F-statistic.

***Example* 2.** The conditional probability that the variate X should lie within the range dx for a given σ is given by

$$(\sigma\sqrt{2\pi})^{-1}\exp\{-\tfrac{1}{2}(x-\mu)^2/\sigma^2\}\,dx,\ -\infty<x<\infty$$

while the probability of σ itself lying within the range $d\sigma$ is

$$\sigma_0^{-2}\exp\{-\sigma^2/2\sigma_0^2\}\,\sigma\,d\sigma,\ 0<\sigma<\infty. \qquad [\sigma_0=\text{const.}]$$

Show that the unconditional (i.e. marginal) distribution of X has the p.d.f.

$$(2\sigma_0)^{-1}\exp\{-|x-\mu|/\sigma_0\},\ -\infty<x<\infty.$$

Solution. We are given that $f(x\mid\sigma)$ is $N(\mu,\sigma^2)$ and $f(\sigma)=(\sigma/\sigma_0^2)\exp\{-\sigma^2/2\sigma_0\},\ 0<\sigma<\infty$

$$\therefore\quad \mathrm{E}\{e^{tX}\}=\mathrm{E}\{\mathrm{E}(e^{tX}\mid\sigma)\}=\mathrm{E}\{e^{\mu t+\frac{1}{2}\sigma^2t^2}\}. \qquad [\because\ X\sim N(\mu,\sigma^2)]\ \text{Double-E Rule}$$

$$=e^{\mu t}\mathrm{E}_\sigma\{e^{\sigma^2t^2/2}\}=e^{\mu t}\int_0^\infty\frac{\sigma}{\sigma_0^2}e^{-\sigma^2/2\sigma_0^2}e^{\sigma^2t^2/2}\,d\sigma=\frac{e^{\mu t}}{\sigma_0^2}\int_0^\infty\sigma\,e^{-(\lambda^2-t^2)\sigma^2/2}\,d\sigma,\ [\lambda^2=1/\sigma_0^2]$$

$$=\frac{\lambda^2e^{\mu t}}{(\lambda^2-t^2)}\int_0^\infty e^{-z}\,dz,\ \left[z=\tfrac{1}{2}(\lambda^2-t^2)\sigma^2,\ \sigma\,d\sigma=\frac{dz}{\lambda^2-t^2}\right]$$

$$=\frac{\lambda^2e^{\mu t}}{\lambda^2-t^2}.$$

This is the m.g.f. of Lap (λ,μ) distribution. Hence the reqd. p.d.f. is given by

$$f(x)=\frac{1}{2}\lambda e^{-\lambda|x-\mu|}=\frac{1}{2\sigma_0}e^{-|x-\mu|/\sigma_0},\ -\infty<x<\infty.$$

***Example* 3.** Let $X\sim U(0,1)$ and $(Y\mid X=x)\sim$ bin (n,x), i.e.

$$P\{Y=y\mid X=x\}={}^nC_y\,x^y(1-x)^{n-y},\ y=0,1,\ldots,n.$$

Find the distribution of Y. What is $\mathrm{E}(Y)$?

Solution. $\quad \mathrm{E}(Y)=\mathrm{E}_X[\mathrm{E}_Y(Y\mid X)]=\mathrm{E}_X(nX)=n\mathrm{E}(X)=n(\tfrac{1}{2})=\tfrac{1}{2}n.$

$$p=P\{Y=y\}=\int_0^1P(Y=y\mid X=x)\,.\,f(x)\,dx \qquad [\text{Multi-Stage Rule}]$$

$$=\int_0^1\binom{n}{y}x^y(1-x)^{n-y}\,.\,1\,dx=\binom{n}{y}\int_0^1x^{(y+1)-1}(1-x)^{(n-y+1)-1}\,dx$$

$$=\binom{n}{y}B(y+1,n-y+1)=\binom{n}{y}\frac{\Gamma(y+1)\,\Gamma(n-y+1)}{\Gamma(n+2)}=\frac{n!\,y!\,(n-y)!}{(n-y)!\,y!\,(n+1)!}$$

$$=1/(n+1).$$

Thus, $P\{Y=y\}=1/(n+1),\ 0\le y\le n$ i.e. $Y\sim U\{0,1,2,\ldots,n\}$. This also gives $\mathrm{E}(Y)=\tfrac{1}{2}n$.

***Example* 4.** The distribution of X given θ is Unif$(0,\theta)$. The prior distribution of θ has p.d.f. $g(\theta)=k/\theta^2,\ \theta>k$. Find the posterior distribution of θ.

Solution. Here $\quad f(x\mid\theta)=1/\theta,\ 0\le x\le\theta$ and $g(\theta)=k/\theta^2,\ \theta>k$

$\therefore\quad f(x,\theta)=f(x\mid\theta)\,g(\theta)=k/\theta^3,\ 0\le x\le\theta,\ \theta>k.$

The marginal density $f(x)$ depends on whether $x < k$ or $x > k$. Thus :

If $x < k < \theta$, then $f_1(x) = \int_k^\infty f(x, \theta)\, d\theta = \left[\dfrac{k}{2\theta^2}\right]_\infty^k = \dfrac{1}{2k}$.

If $k < x < \theta$, then $f_1(x) = \int_x^\infty \dfrac{k}{\theta^3}\, d\theta = \dfrac{k}{2x^2}$, $k < x < \theta < \infty$.

There is *knock-on* effect on the form of $f(\theta \mid x)$. Thus

$$f(\theta \mid x) = \frac{k/\theta^3}{1/2k} = \frac{2k^2}{\theta^3}, \quad x < k < \theta; \quad f(\theta \mid x) = \frac{k/\theta^3}{k/2x^2} = \frac{2x^2}{\theta^3}, \ \theta > x.$$

***Example* 5.** Let X be bin $(n, 1/2)$ and let the compounding distribution of the parameter n be $f(n) = n/3$, $n = 1, 2$. Find the unconditional distribution of X and its variance.

Solution. Given : $f(x \mid n) = \binom{n}{x}\left(\frac{1}{2}\right)^n$, $x = 0, 1, 2, ..., n$; $P(N = n) = \frac{1}{3}n$, $n = 1, 2$,

$$P(X = x, N = n) = P(X = x \mid N = n)\, P(N = n) = \tfrac{n}{3}\binom{n}{x}\left(\tfrac{1}{2}\right)^n, \ x = 0, 1, 2, ..., n, \ n = 1, 2$$

By the Theorem of Total Probability : $P(X = x) = \sum_n P(X = x, N = n)$

$$P(X = 0) = P(X = 0, N = 1) + P(X = 0, N = 2) = \left(\tfrac{1}{6}\right) + \left(\tfrac{1}{6}\right) = \tfrac{1}{3}$$

$$P(X = 1) = P(X = 1, N = 1) + P(X = 1, N = 2) = \left(\tfrac{1}{6}\right) + \left(\tfrac{1}{3}\right) = \tfrac{1}{2}$$

$$P(X = 2) = P(X = 2, N = 2) = \tfrac{1}{6} \qquad (n \geq x)$$

Thus the unconditional distribution of X is

$$P(X = 0) = \tfrac{1}{3}, \ P(X = 1) = \tfrac{1}{2}, \ P(X = 2) = \tfrac{1}{6}$$

$$E(X) = \Sigma x\, P(X = x) = \tfrac{1}{2} + 2 . \tfrac{1}{6} = \tfrac{5}{6}; \ E(X^2) = \Sigma x^2 P(X = x) = 1 . \tfrac{1}{2} + 4 . \tfrac{1}{6} = \tfrac{7}{6}.$$

$$\text{Var}(X) = \left(\tfrac{7}{6}\right) - \left(\tfrac{25}{36}\right) = \left(\tfrac{17}{36}\right).$$

***Example* 6.** If $X \mid Y = y$ is gam (a, Y) and Y is gam $(b, 1)$, show that the unconditional distribution of X is $B_{II}(a, b)$. Further, if $Z \mid (X = x, Y = y)$ is bin $(n, x/(1 + x))$, find the unconditional distribution of Z and its variance.

Solution. (i) From $X \mid y \sim$ gam (a, y) we instantly get $f(x \mid y) = e^{-yx} x^{a-1} y^a/\Gamma(a)$.

$$f(x, y) = f(x \mid y)\, f_2(y) = \frac{y^a x^{a-1} e^{-yx}}{\Gamma(a)} \cdot \frac{y^{b-1} e^{-y}}{\Gamma(b)} = \frac{e^{-y(1+x)} y^{a+b-1} x^{a-1}}{\Gamma(a)\,\Gamma(b)}. \qquad ...(i)$$

To find the distribution of X, we integrate **out** y from (i); thus

$$f_1(x) = \frac{x^{a-1}}{\Gamma(a)\,\Gamma(b)} \int_0^\infty y^{a+b-1} e^{-y(1+x)} dy, \quad x \geq 0$$

$$= \frac{x^{a-1}}{\Gamma(a)\,\Gamma(b)} \cdot \frac{\Gamma(a+b)}{(1+x)^{a-b}} = \frac{x^{a-1}}{B(a, b)\,(1+x)^{a+b}}, \ x \geq 0.$$

This shows that the unconditional distribution of X is $B_{II}(a, b)$.

(ii) Since $Z \mid (X = x, Y = y)$ is bin $(n, x/(1 + x))$, we have

$$f(z \mid x, y) = \binom{n}{z}\left(\frac{x}{1+x}\right)^z\left(1-\frac{x}{1+x}\right)^{n-z} = \binom{n}{z}\frac{x^z}{(1+x)^n}.$$

$$\therefore \quad g(z, x, y) = f(z \mid x, y) \,.\, f(x, y) \quad \text{(joint p.m.f.)}$$

$$= \binom{n}{z}\frac{x^z}{(1+x)^n}\cdot\frac{y^{a+b-1}\, e^{-y(1+x)}\, x^{a-1}}{\Gamma(a)\,\Gamma(a)}. \qquad [\text{by (i)}],\ x > 0,\ y > 0.$$

To find out the distribution of Z, we integrate **out** x and y from above.

$$g(z) = \binom{n}{z}\int_0^\infty\int_0^\infty \frac{y^{a+b-1}\, e^{-y(1+x)}\, x^{a+z-1}}{\Gamma(a)\,\Gamma(b)\,(1+x)^n}\,dx\,dy = \binom{n}{z}\int_0^\infty \frac{\Gamma(a+b)}{\Gamma(a)\,\Gamma(b)}\,\frac{x^{a+z-1}\,dx}{(1+x)^n\,.\,(1+x)^{a+b}}$$

$$= \binom{n}{z}\frac{1}{B(a,b)}\int_0^\infty \frac{x^{a+z-1}\,dx}{(1+x)^{(a+z)+(b+n-z)}} = \binom{n}{z}\frac{B(a+z, n+b-z)}{B(a,b)},\quad z = 0, 1, 2, \ldots, n.$$

This is unconditioned distribution of Z. Now

$$\mathrm{E}(Z^{(k)}) = \frac{1}{B(a,b)}\sum_{z=0}^{n} z^{(k)}\binom{n}{z} B(a+z, n+b-z)$$

$$= \frac{n^{(k)}}{B(a,b)}\sum_{z=k}^{n}\binom{n-k}{z-k}\int_0^1 t^{a+z-1}\,(1-t)^{n+b-z-1}\,dt,$$

$$= \int_0^1 t^{k+a-1}\,(1-t)^{b-1}dt\,\sum_{z=k}^{0}\binom{n-k}{z-k}t^{z-k}\,(1-t)^{n-z}, \qquad \left[\lambda = \frac{n^{(k)}}{B(a,b)}\right]$$

$$= \lambda\int_0^1 t^{k+a-1}(1-t)^{b-1}[(1-t)+t]^{n-k}dt = \lambda\int_0^1 t^{k+a-1}(1-t)^{b-1}\,dt = \lambda B(k+a, b)$$

$$= n^{(k)}\,\frac{\Gamma(k+a)\,\Gamma(b+a)}{\Gamma(a)\,\Gamma(k+a+b)} = n^{(k)}\,.\,\frac{a^{[k]}}{(a+b)^{[k]}} \qquad [\text{Putting for } \lambda]$$

$$\therefore \quad \mathrm{E}(Z) = \frac{na}{a+b},\quad \mathrm{E}(Z^{(2)}) = \frac{n(n-1)\,a(a+1)}{(a+b)\,(a+b+1)}.$$

Var $(Z) = \mathrm{E}(Z^{(2)}) + \mathrm{E}(Z) - \mathrm{E}^2(Z) = nab\,(n + b + a)/(a + b)^2\,(a + b + 1)$.

Exercises

1. The distribution of Y is $P(Y = n) = (\frac{1}{2})^n$, $n = 1, 2, 3, \ldots$ The conditional distribution of the continuous variate X for a discrete variate Y assuming a value n is
$$dF = n(1-x)^{n-1}\,dx,\ 0 \le x \le 1.$$
Find the marginal distribution of X.
2. Given $f(x \mid y) = e^{-y}\, y^x/x\,!$ and $h(y) = e^{-y}$, where Y is continuous $y \ge 0$ and X is discrete, i.e. $x = 0$, 1, 2, ... Show that the marginal distribution of X is geometric, i.e. $g(x) = (1/2)^{x+1}$. Find also the conditional means of X and Y.
3. Let X be bin (n, p) and let the compounding distribution of the parameter n be bin (m, p_1). Show that the unconditional distribution of X is bin (m, pp_1).

4. Let X be bin (n, p) and let the compounding distribution of the parameter p be $\beta_1(a, b)$. Show that the unconditional distribution of X is

$$f(x) = B(a + x, b + n - x)\,[(n + 1)\,B(a, b)\,B(x + 1, n - x + 1)]^{-1} \quad x = 0, 1, 2, \ldots, n$$

Hence find its mean and variance. Find the corresponding quantities if p is $U(0, 1)$.

5. *Neyman's Contagious Compound distn.* Let X be Pois (λ, y) where y itself is an observation of a variate $Y \sim$ Pois (m). Find the unconditional distribution of X and show that its mean is less than its variance.

6. Let $X \sim$ Expo (Y) where Y itself is distributed as gam (r, λ). Find the unconditional density of X and the conditional density of Y given $X = x$.

7. Let X be Pois (λ) and let the compounding distribution for λ possess a Ch. Function $\phi(t)$. Show that the unconditional distribution of X has the Ch. Function $\phi[i(1 - e^{it})]$ and that its mean and variance are λ and $E(\lambda) + \text{Var}(\lambda)$ respectively.

8. Let $X \sim$ Expo (Y) and let the compounding distribution of the parameter Y be a discrete variate with p.m.f. $P(Y = 1) = 1/2$, $P(Y = 2) = \frac{1}{3}$, $P(Y = 3) = \frac{1}{6}$.

   Find the unconditional distribution of X and its mean and variance. Find also $P(X \geq 2)$. Find the probability that Y is an odd number and X is atmost 2.

9. Let $\langle X_n \rangle$ be *a* sequence of i.i.d. LD (θ) variates [ℓn=series distribution with parameter θ]. If $S = X_1 + X_2 + \ldots + X_N$ where N is Pois (λ) show that S is distributed as *a* negative binomial variate provided $\lambda/\ell n\,(1 - \theta)$ is an integer.

10. Let X be $U(0, Y)$ and let the compounding distribution of Y be again $U(0, 1)$. Find the unconditional distribution of X. Show that $E(Y \mid x) \leq 1$ and hence deduce that $x - 1 \geq \ell n\, x$ for $0 < x \leq 1$. If the compounding distribution of the parameter Y is Expo (λ), find $E(X)$, Var (X), $E(Y \mid x)$ and $E(X \mid y)$.

11. Let $\langle X_k \rangle$ be a sequence of Ber (p). Show that $(\sum_{1}^{N} X_k) \sim$ Pois (λp) when $N \sim$ Pois (λ).

***Total understanding makes one very indulgent.* (Madame De Stael)**

Imagination is more powerful than knowledge. (Albert Einstein)

Bivariate Normal Distribution

20

20-10. *Bivariate Normal Distribution, Marginal and Conditional Distributions*

Definition. The r.v.s X and Y are said to possess a Bivariate Normal Distribution with five finite parameters $\mu_1, \mu_2, \sigma_1 > 0, \sigma_2 > 0$; ρ, if their joint p.d.f. is given by

$$f(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-\frac{1}{2}Q(x,y)}, \quad -\infty < x, y < \infty, \, |\rho| < 1. \qquad ...(1)$$

where $Q(x, y) = \dfrac{1}{(1-\rho^2)}\left\{\left(\dfrac{x-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\dfrac{x-\mu_1}{\sigma_1}\right)\left(\dfrac{y-\mu_2}{\sigma_2}\right) + \left(\dfrac{y-\mu_2}{\sigma_2}\right)^2\right\}$

Marginal Distributions. It is convenient to write

$$u = \frac{x-\mu_1}{\sigma_1}, \; v = \frac{y-\mu_2}{\sigma_2}; \; w = \frac{(v-\rho u)}{\sqrt{1-\rho^2}}, \qquad \left(\text{or } w = \frac{u-\rho v}{\sqrt{1-\rho^2}}\right)$$

Then $Q(x, y) = u^2 + \dfrac{(v-\rho u)^2}{1-\rho^2} = v^2 + \dfrac{(u-\rho v)^2}{1-\rho^2}$

Take : $Q(x, y) = u^2 + \dfrac{(v-\rho u)^2}{1-\rho^2} = u^2 + \left\{\dfrac{y_2 - [\mu_2 + \rho\sigma_2(x-\mu_1)/\sigma_1]}{\sigma_2\sqrt{1-\rho^2}}\right\}^2$

$$\therefore \qquad Q(x, y) = u^2 + [(y-m_2)/\sigma_2']^2 \qquad ...(2)$$

where $\qquad m_2 = \mu_2 + \rho\sigma_2(x-\mu_1)/\sigma_1; \;\; \sigma_2'^2 = \sigma_2^2(1-\rho^2) \qquad ...(3)$

From (1) and (2),

$$f(x, y) = \left\{\frac{e^{-u^2/2}}{\sqrt{2\pi}\,\sigma_1}\right\}\left\{\frac{e^{-(y-m_2)^2/2\sigma_2'^2}}{\sqrt{2\pi}\,\sigma_2'}\right\} \qquad ...(4)$$

$$\therefore \qquad f_1(x) = \sum_y f(x, y) = \frac{e^{-u^2/2}}{\sqrt{2\pi}\sigma_1}\left\{\int_{-\infty}^{\infty} \frac{e^{-(y-m_2)^2/2\sigma_2'^2}}{\sqrt{2\pi}\,\sigma_2'}\partial y\right\}$$

The integral, being area under $N(m_2, \sigma_2'^2)$ is unity.

$$\therefore \qquad f_1(x) = \frac{e^{-u^2/2}}{\sqrt{2\pi}\,\sigma_1} = \frac{e^{-(x-\mu_1)^2/2\sigma_1^2}}{\sqrt{2\pi}\,\sigma_1}, \quad -\infty < x < \infty. \qquad ...(5)$$

Incidently, $\sum_x \sum_y f(x, y) = \sum_x f_1(x) = 1.$

Thus (1) is a bonafide p.d.f. as it is certainly non-negative. From (5) and hence interchange of X, Y, we conclude that

$$X \sim N(\mu_1, \sigma_1^2),\ Y \sim N(\mu_2, \sigma_2^2).$$

Conditional densities :

From Eq. (4), we readily observe that

$$f(x, y) = f_1(x) \,.\, f(y \mid x)$$

$$\therefore \qquad (Y \mid X = x) \sim N(m_2, \sigma'^2_2)$$

where $m_2 = \mathrm{E}(Y \mid x) = \mu_2 + \rho \dfrac{\sigma_2}{\sigma_1}(x - \mu_1)$; $\quad \mathrm{Var}\,(Y \mid x) = \sigma_2'^2 = \sigma_2^2\,(1 - \rho^2)$.

Similarly, $(X \mid y) \sim N(m_1, \sigma_1'^2)$, $\quad m_1 = \mu_1 + \rho \dfrac{\sigma_1}{\sigma_2}(y - \mu_2)$, $\quad \sigma_1'^2 = \sigma_1^2(1 - \rho^2)$.

Note. We write

$$K = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix},\ \mathbf{x} = (x, y), \boldsymbol{\mu} = (\mu_1, \mu_2).\ |K| = \sigma_1^2\sigma_2^2(1 - \rho^2), \text{ then}$$

$$f(x, y) = \frac{1}{(\sqrt{2}\pi)^2 \sqrt{|K|}} \,.\exp\left[-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T K^{-1}(\mathbf{x} - \boldsymbol{\mu})\right],\ -\infty < \mathbf{x} < \infty$$

This expression can be readily extended to n variables $X_1, X_2, \ldots, X_n$.

An alternate modern definition. A random vector (X, Y) is said to possess bivariate normal distribution, if for **all** a and b, $Z = aX + bY$ is a normal variate.

Remark. The equivalence of two definitions is shown in §20-50.

It follows that $f(x, y)$ defines a continuous bivariate distribution. The variates X and Y with joint p.d.f. $f(x, y)$ given by (2) are said to have a bivariate normal distribution with five parameters : $\mu_X, \mu_Y, \sigma_X, \sigma_Y$. When $\sigma_X = \sigma_Y = 1$ and $\mu_X = \mu_Y = 0$, then X and Y are said to have standard bivariate normal distribution.

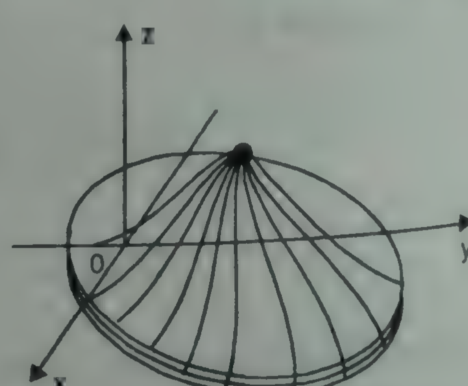
Notation. A bivariate normal distribution with five parameters $\mu_X(=\mu_1)$, $\mu_Y(=\mu_2)$, $\sigma_X^2(=\sigma_1^2)$, $\sigma_Y^2(=\sigma_2^2)$ and ρ is denoted by

$$(X, Y) \sim \mathrm{BVN}\ (\mu_1, \mu_2, \sigma_1^2, \sigma_2^2; \rho)$$

Remarks. If $\rho = \pm 1$, the density function in (2) is undefined, indeed the distribution is then called *singular*, with the unit of probability distribution along a st. line $y = mx + c$. In this case, we may argue that X and Y are still marginally normal, but linearly related with probability 1, that is

$$P\{Y = \mu_2 + (\sigma_2/\sigma_1)\,(X - \mu_1)\} = 1.$$

The adjoining figure shows the BVN-surface.



Bivariate Normal Distribution

20-20. *Moment Generating Function*

Recall : $X \sim N(\mu_1, \sigma_1^2)$, $\quad (Y \mid x) \sim N(m_2, \sigma_2'^2)$

where $m_2 = \mu_2 + \rho(\sigma_2/\sigma_1)(x - \mu_1)$, $\quad \sigma_2'^2 = \sigma_2^2(1-\rho^2)$.

We evaluate m.g.f. by Double-E Rule. Note that m_2 contains x. Now

$$\begin{aligned} M(t_1, t_2) &= \mathrm{E}(e^{t_1X + t_2Y}) = \mathrm{E}\{\mathrm{E}[e^{t_1X+t_2Y} \mid X]\} \\ &= \mathrm{E}\{e^{t_1X}\,\mathrm{E}(e^{t_2Y} \mid X)\}, \qquad [\text{use m.g.f. of } (Y \mid x) \sim N(m_2, \sigma_2'^2)] \\ &= \mathrm{E}\{e^{t_1X} . e^{m_2t_2 + \frac{1}{2}\sigma_2'^2t_2^2}\} \\ &= \exp\{\mu_2t_2 - \rho(\sigma_2/\sigma_1)\mu_1t_2 + \tfrac{1}{2}\sigma_2'^2t_2^2\}\{\mathrm{E}e^{[t_1 + \rho(\sigma_2/\sigma_1)t_2]X}\}, \quad [\text{use m.g.f. of } X] \end{aligned}$$

$$= \exp\{\mu_2t_2 + \tfrac{1}{2}\sigma_2^2t_2^2 - \tfrac{1}{2}\rho^2\sigma_2^2t_2^2 - \rho(\sigma_2/\sigma_1)\mu_1t_2\}\{\exp \mu_1(t_1 + \rho(\sigma_2/\sigma_1)t_2 + \tfrac{1}{2}\sigma_1^2[t_1 + \rho(\sigma_2/\sigma_1)t_2]^2\}$$

This instantly provides

$$\boxed{M(t_1, t_2) = \exp\{\mu_1t_1 + \mu_2t_2 + \tfrac{1}{2}(\sigma_1^2t_1^2 + 2\rho\sigma_1\sigma_2t_1t_2 + \sigma_2^2t_2^2)\}}$$

Cor. Let $U = \left(\dfrac{X-\mu_1}{\sigma_1}\right)$, $V = \left(\dfrac{Y-\mu_2}{\sigma_2}\right)$, $W = \dfrac{V - \rho U}{\sqrt{1-\rho^2}}$

Then $U \sim N(0,1)$ and $W \sim N(0,1)$ are independent.

Proof.

$$\begin{aligned} M(t_1, t_2) &= \mathrm{E}\{e^{t_1U + t_2W}\} = \mathrm{E}\{e^{t_1U + t_2(V-\rho U)/\sqrt{1-\rho^2}}\} \\ &= \mathrm{E}\{e^{t_3U + t_4V}\},\ t_3 = t_1 - [\rho t_2/\sqrt{1-\rho^2}],\ t_4 = t_2/\sqrt{1-\rho^2} \end{aligned}$$

This expectation is the m.g.f. of BVN (0, 0, 1, 1; ρ). Hence by above

$$\begin{aligned} M(t_1, t_2) &= e^{\frac{1}{2}(t_3^2 + 2\rho t_3t_4 + t_4^2)} = e^{\frac{1}{2}(t_1^2 + t_2^2)} \\ &= e^{\frac{1}{2}t_1^2} . e^{\frac{1}{2}t_2^2}\ [= M_U(t_1)\,M_W(t_2)] \end{aligned}$$

Thus, $U \sim N(0,1)$ and $W \sim N(0,1)$ are independent.

20-21. *Five Parameters Interpreted*

Maclaurin's series for the exponential function of the m.g.f. of bivariate normal is

$$\begin{aligned} M(t_1, t_2) &= 1 + [\mu_1t_1 + \mu_2t_1 + \tfrac{1}{2}(\sigma_1^2t_1^2 + 2\rho\sigma_1\sigma_2t_1t_2 + \sigma_2^2t_2^2)] \\ &\quad + \tfrac{1}{2}[\mu_1t_1 + \mu_2t_2 + \tfrac{1}{2}(\sigma_1^2t_1^2 + 2\rho\sigma_1\sigma_2t_1t_2 + \sigma_2^2t_2^2)]^2 + \ldots \qquad \ldots(1) \end{aligned}$$

$\because \quad \mathrm{E}(X^rY^s) = \text{coeff. of } (t_1^r\,t_2^s/r!s!) \text{ in } M(t_1, t_2)$

$\therefore\ \mathrm{E}(X) = \mu_1,\ \mathrm{E}(Y) = \mu_2,\ \mathrm{E}(X^2) = \sigma_1^2 + \mu_1^2,\ \mathrm{E}(Y^2) = \sigma_2^2 + \mu_2^2,\ \mathrm{E}(XY) = \rho\sigma_1\sigma_2 + \mu_1\mu_2.$

These instantly provide means and variances [Var $X = \mathrm{E}(X^2) - \mathrm{E}^2(X)$, etc.]

$$\mathrm{Var}(X) = \sigma_1^2,\ \mathrm{Var}(Y) = \sigma_2^2,\ \mathrm{Cov}(X,Y) = \rho\,\sigma_1\sigma_2,\ \text{i.e. } \mathrm{Corr}(X,Y) = \rho.$$

Thus, all the five parameters stand interpreted.

20-30. *Regression Function and Regression Curve*

If X, Y are jointly distributed and $E(Y \mid x)$ exists $\forall$ real x, then $r(x) = E(Y \mid x)$ is called the *regression function* of Y on X. The graph of $r(x)$ is called the *regression curve* of Y on X. For the BVN distribution

$$(Y \mid x) \sim N[\mu_2 + \rho(\sigma_2/\sigma_1)(x - \mu_1),\ \sigma_2^2(1 - \rho^2)].$$

$$(X \mid y) \sim N[\mu_1 + \rho(\sigma_1/\sigma_2)(y - \mu_2),\ \sigma_1^2(1 - \rho^2)].$$

It follows that the regression functions are linear :

$$E(Y \mid x) = \mu_2 + \rho(\sigma_2/\sigma_1)(x - \mu_1); \qquad E(X \mid y) = \mu_1 + \rho(\sigma_1/\sigma_2)(y - \mu_2).$$

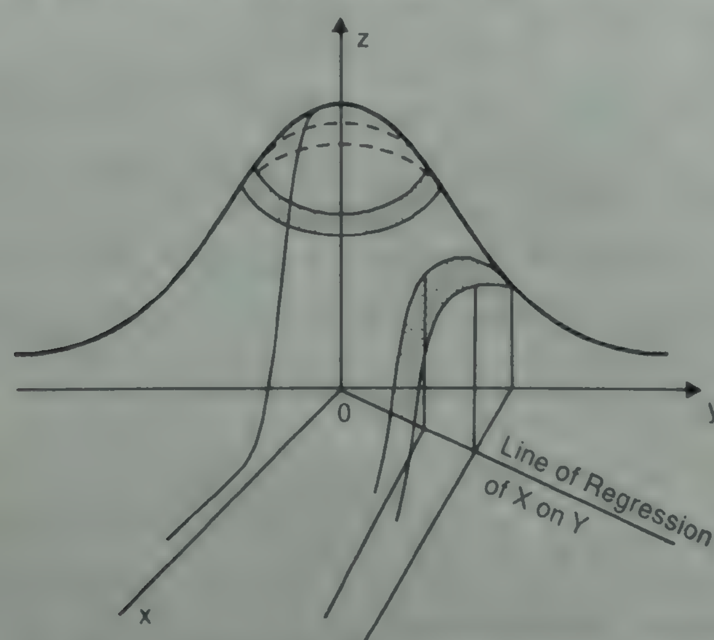
Obviously, the regression curves are st. lines. Further, conditional variates are constant. The linear regression of X on Y (or Y on X) is homoscedastic (i.e. has equal scattering).

Comments. All results relating to linear regressions hold true for BVN variates.

20-40. *Level Curves of Normal Correlation Surface*

Here : $f(x, y) = (2\pi\sigma_1\sigma_1\sqrt{1-\rho^2})^{-1} \exp(-Q/2),\ -\infty < x,\ y < \infty.$

with $Q = \{[(x - \mu_1)/\sigma_1]^2 - 2\rho[(x - \mu_1)(y - \mu_2)/\sigma_1\sigma_2] + [(y_2 - \mu_2)/\sigma_2]^2\}/(1 - \rho^2).$



BIVARIATE NORMAL DISTRIBUTION

The surface $z = f(x, y)$ is called the Normal correlation surface or bivariate normal density surface. Level curves are the contours at fixed *elevations* of this surface or as the intersection of this surface with planes parallel to the x-y plane.

The curves along which the p.d.f. is constant are the homothetic ellipses

$$(x^2/\sigma_1^2) - 2\rho(xy/\sigma_1\sigma_2) + (y^2/\sigma_2^2) = \lambda^2 \qquad \text{...(1)}$$

[$\because$ (coeff. $xy)^2$ – (coeff. x^2) (coeff. y^2) = $4[(\rho^2 - 1)/\sigma_1^2\sigma_1^2 < 0]$ where $\mu_1 = \mu_2 = 0$ for simplicity; with their centres at (0, 0) and major axis inclined at angle θ with x-axis, given by

$$\tan 2\alpha = 2\rho\sigma_1\sigma_2/(\sigma_1^2 - \sigma_2^2).$$

For homothetic ellipses, the line of regression of y on x is conjugate to the axis of y. The line $x = a$ (parallel to y-axis) intercepts the ellipse (1) into two points whose ordinates y_1, y_2 are the roots of the Eq.

$$(y^2/\sigma_2^2) - (2\rho a/\sigma_1/\sigma_2)y + [(a^2/\sigma_1^2) - \lambda^2] = 0.$$

The coordinates of the mid-point of the intersection are $x = a$, $y = \frac{1}{2}(y_1 + y_2) = a\rho\sigma_2/\sigma_1$.

Hence, the locus of the mid-point is $y = (\rho\sigma_2/\sigma_1)x$, which is the line of regression of Y on X. A similar argument shows that the line of regression of X on Y is conjugate to the axis of x.

Problem. Show that if (X, Y) is BVN $(0, 0, \sigma_1^2, \sigma_2^2; 0)$, the point of inflexion of the curve of intersection of the normal correlation surface by planes through the z-axis, lie on the elliptic cylinder $(x^2/\sigma_1^2) + (y^2/\sigma_2^2) = 1$.

Solution. The normal correlation surface is

$$z = f(x, y) = (2\pi\sigma_1\sigma_2)^{-1} \exp\{-[(x^2/\sigma_1^2) + (y^2/\sigma_2^2)]/2\}.$$

Any plane through z-axis is $y = cx$, so that the curve of intersection is

$$z = (2\pi\sigma_1\sigma_2)^{-1} \exp[-\tfrac{1}{2}(\sigma_1^{-2} + c^2\sigma_2^{-2})x^2].$$

$$\therefore \qquad \ln z = K - \tfrac{1}{2}Ax^2 \quad [A = \sigma_1^{-2} + c^2\sigma_2^{-2};\ K = \text{const.}] \qquad \ldots(1)$$

Points of inflexion are given by $z'' = 0$, $z''' \neq 0$. Hence differentiating (1)

$$(z'/z) = -Ax, \qquad (z''/z) - (z'/z)^2 = -A. \qquad \ldots(2)$$

Putting $z'' = 0$, and eliminating (z'/z), we get $A^2x^2 = A$. This gives $x = (A)^{-1/2}$. Thus the points of inflexion are $x = A^{-1/2}$, $y = cA^{-1/2}$, $(y = cx)$.

For the locus we need eliminate $c(= y/x)$. Thus

$$x^2 = [(1/\sigma_1^2) + (c^2/\sigma_2^2)]^{-1} = [(1/\sigma_1^2) + (y^2/x^2\sigma_2^2)]^{-1} = \sigma_1^2\sigma_2^2x^2/(\sigma_1^2y^2 + \sigma_2^2x^2)$$

$$\therefore \quad \sigma_2^2x^2 + \sigma_1^2y^2 = \sigma_1^2\sigma_2^2 \Rightarrow (x^2/\sigma_1^2) + (y^2/\sigma_2^2) = 1. \qquad \text{(Elliptic Cylinder)}$$

20-50. *Independence versus Uncorrelatedness*

Let (X, Y) be BVN $(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2; \rho)$. Then X and Y are stochastically independent iff $\rho = 0$.

Proof. Recall : $M(t_1, t_2) = \exp[\mu_1 t_1 + \mu_2 t_2 + \frac{1}{2}(\sigma_1^2 t_1^2 + 2\rho\sigma_1\sigma_2 t_1 t_2 + \sigma_2^2 t_2^2)]$.

(i) Let X and Y be independent, then Cov $(X, Y) = 0$, i.e. $\rho = 0$. [which is always true].

(ii) Let $\rho = 0$, i.e. X and Y are correlated, then

$$M(t_1, t_2) = \exp[t_1\mu_1 + t_2\mu_2 + \tfrac{1}{2}(\sigma_1^2 t_1^2 + \sigma_2^2 t_2^2)] = \exp[\mu_1 t_1 + \tfrac{1}{2}\sigma_1^2 t_1^2].\exp[\mu_2 t_2 + \tfrac{1}{2}\sigma_2^2 t_2^2] = M_X(t_1).M_Y(t_2).$$

This shows that X and Y are independent Normal variates.

Note. Two uncorrelated variates that have a BVN distribution are always independent.

20-51. *Theorem. Independence of Sample Mean and Sample Variance*

Let X_i, $1 \le i \le n$ be indep. $N(\mu, \sigma^2)$ variates. Then, $\overline{X}_n$ and $S^2 = \Sigma(X_i - \overline{X})^2/(n-1)$ are independent.

Proof. The variates $X_1 - \overline{X}_n, X_2 - \overline{X}_n, \ldots, X_n - \overline{X}_n$ are jointly normal. Further more

$\mathrm{Cov}(X_i, \overline{X}_n) = (1/n)[\mathrm{Cov}(X_i, X_1) + \ldots + \mathrm{Cov}(X_i, X_n)] = \sigma^2/n$. $[\mathrm{Cov}(X_i, X_i) = \sigma^2, \mathrm{Cov}(X_i, X_j) = 0, j \ne i$.

$$\therefore \quad \mathrm{Cov}(X_i - \overline{X}_n, \overline{X}_n) = \mathrm{Cov}(X_i, \overline{X}_n) - \mathrm{Cov}(X_n, \overline{X}_n) = (\sigma^2/n) - (\sigma^2/n) = 0.$$

It follows that $X_1 - \overline{X}_n, \ldots, X_n - \overline{X}_n$ are uncorrelated to X_n and hence independent (joint normal distribution). Thus $\Sigma(X_i - \overline{X}_n)^2$ and $\overline{X}_n$ are also independent. This suffices to provide the result.

Comments. Without assuming joint normality of $X_i - \overline{X}_n$, the result has been established in §16-54.

20-52. *Rotation for Securing Independence of Variates*

Let $(X, Y) \sim \mathrm{BVN}(0, 0, \sigma_1^2, \sigma_2^2; \rho)$ be a point in the XOY plane. Rotating axes through θ, the point (X, Y) becomes (U, V) w.r.t. new axes. These are related by

$$\left.\begin{aligned} U &= X\cos\theta + Y\sin\theta \\ V &= -X\sin\theta + Y\cos\theta \end{aligned}\right\} \text{ or } \begin{bmatrix} U \\ V \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} X \\ Y \end{bmatrix} \text{ or } \begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} U \\ V \end{bmatrix}.$$

$$\mathrm{Cov}(U, V) = \mathrm{E}(UV) = \mathrm{E}[(Y^2 - X^2)\sin\theta\cos\theta + XY(\cos^2\theta - \sin^2\theta)]$$

$$\sigma_{UV} = 0, \Rightarrow \{\mathrm{E}(Y^2) - \mathrm{E}(X^2)\}(\sin 2\theta/2) + \mathrm{E}(XY)\cos 2\theta = 0.$$

$$\therefore \quad \tan 2\theta = 2\rho\,\sigma_1\,\sigma_1/(\sigma_1^2 - \sigma_2^2) \text{ or } \theta = (\tfrac{1}{2})\tan^{-1}[2\rho\sigma_1\sigma_2/(\sigma_1^2 - \sigma_2^2)]. \qquad \ldots(1)$$

Note. In obtaining (1), nature of BVN is not used. Hence (1) is valid for arbitrary variates.

***Example* 1.** Let $(X, Y) \sim \mathrm{BVN}(0, 0, 9, 4; \rho\ ?)$. If coordinate rotation through $\alpha = -\pi/8$, results in uncorrelated variates (U, V), find ρ.

Solution. $$\rho = \frac{\sigma_1^2 - \sigma_2^2}{2\sigma_1\sigma_2}\tan 2\theta = \frac{9-4}{2.3.2}\tan\left(-\frac{\pi}{4}\right) = -\frac{5}{12}.$$

***Example* 2.** Let $(X, Y) \sim \mathrm{BVN}(0, 0, \sigma^2, \sigma^2; -1)$. Find the transformation matrix such that $(X, Y) \to (U, V)$ independent variates.

Solution. $$\tan 2\theta = \frac{2\rho\sigma_1\sigma_2}{\sigma_1^2 - \sigma_2^2} = \frac{2(-1)\sigma^2}{\sigma_1^2 - \sigma_2^2} \to -\infty \Rightarrow 2\theta = -\pi/2 \text{ or } \theta = -\pi/4.$$

$$R = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}.$$

20-60. Bevy of Solved Examples

Example 1. If X and Y are i.i.d. $N(0, 1)$ variates, show that the m.g.f. of XY is $(1-t^2)^{-1/2}, -1<t<1$.

Solution.
$$M(t:XY) = E(e^{tXY}) = E\{E[e^{tXY} \mid Y]\} = E\{E[e^{(tY)X}] \mid Y\}, \qquad [\text{use } M(t:X)]$$
$$= E\{e^{\frac{1}{2}t^2Y^2}\}, \qquad [Y^2 \sim \text{gam}\,(1, \tfrac{1}{2}) = \chi^2_{(1)} \text{ with m.g.f.} = (1-2t')^{-\frac{1}{2}}$$
$$= (1-t^2)^{-\frac{1}{2}}, \ (t^2 \neq 1).$$

Important Note. M.G.F. of $Z = XY$, when $(X, Y) \sim \text{BVN}\,(0, 0, \sigma_1^2, \sigma_2^2; \rho)$

Here $\mu_{X \mid y} = m_1 = \rho(\sigma_1/\sigma_2)y$, $\text{Var}\,(X \mid y) = \sigma_1^2(1-\rho^2) = \sigma_1'^2$, $(X \mid y) \sim N(m_1, \sigma_1'^2)$

$$M(t:XY) = E\{e^{tXY}\} = E\{E(e^{tXY} \mid y)\} = E\{E[e^{(ty)X} \mid Y]\}$$
$$= E\{e^{M_1tY + \frac{1}{2}\sigma_1'^2(tY)^2}\}, \qquad [M_1 = \rho(\sigma_1/\sigma_2)Y,\ W = Y/\sigma_2 \sim N(0, 1)]$$
$$= E\{\exp\,(\rho\frac{\sigma_1}{\sigma_2} + \tfrac{1}{2}t^2\sigma_1'^2)\,\sigma_2^2W^2\}, \quad W^2 \sim \chi^2_{(1)} = \text{gam}\,(1, \tfrac{1}{2}),$$
$$= E(e^{TW^2}) = (1-2T)^{-1/2} = [1 - 2\rho\sigma_1\sigma_2 t - \sigma_1^2\sigma_2^2(1-\rho^2)\,t^2]^{-1/2}.$$

Example 2. Sheppard's Result. If (X, Y) is BVN $(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2; \rho)$, show that
$$P\{X > E(X), Y > E(Y)\} = \frac{1}{4} + \frac{1}{2\pi}\sin^{-1}\rho.$$

Solution. $E(X) = \mu_1$, $E(Y) = \mu_2$; put $(x-\mu_1)/\sigma_1 = u$, $(y-\mu_2)/\sigma_2 = v$. Now
$$p = P\{X > \mu_1, Y > \mu_2\} = P\{U > 0, V > 0\} = \frac{1}{2\pi\sqrt{1-\rho^2}}\int_0^\infty\int_0^\infty e^{-(u^2 - 2\rho uv + v^2)/2(1-\rho^2)}\,du\,dv.$$

To evaluate this integral, let $u = r\cos\theta$, $v = r\sin\theta$, $du\,dv = r\,d\theta\,dr$, so
$$p = \frac{1}{2\pi\sqrt{1-\rho^2}}\int_0^{\pi/2} d\theta \int_0^\infty r e^{-r^2(1-2\rho\sin\theta\cos\theta)/2(1-\rho^2)}\,dr$$
$$= \frac{\sqrt{1-\rho^2}}{2\pi}\int_0^{\pi/2}\frac{d\theta}{1-2\rho\sin\theta\cos\theta} = \frac{\sqrt{1-\rho^2}}{2\pi}\int_0^{\pi/2}\frac{\sec^2\theta\,d\theta}{1+\tan^2\theta - 2\rho\tan\theta}$$
$$= \frac{\sqrt{1-\rho^2}}{2\pi}\int_0^\infty\frac{dz}{(z-\rho)^2 + (1-\rho^2)} = \frac{1}{2\pi}\left[\tan^{-1}\left(\frac{z-\rho}{\sqrt{1-\rho^2}}\right)\right]_0^\infty, \qquad [z = \tan\theta,\ \rho^2 < 1]$$
$$= \frac{1}{2\pi}\left\{\frac{\pi}{2} + \tan^{-1}\frac{\rho}{\sqrt{1-\rho^2}}\right\} = \frac{1}{4} + \frac{1}{2\pi}\sin^{-1}(\rho).$$

Remark. The above result follows trivially if we appeal to the integral :
$$\int_0^\infty\int_0^\infty e^{-(ax^2 + 2hxy + by^2)}\,dx\,dy = \frac{1}{2\sqrt{(ab-h^2)}}\cot^{-1}\frac{h}{\sqrt{ab-h^2}}.$$

***Example* 3.** If X is $N(\mu, \sigma^2)$ and $Y \mid x$ is $N(x, \sigma^2)$, show that (X, Y) is BVN $(\mu, \mu, \sigma^2, 2\sigma^2; \rho)$

Solution. $M(t_1, t_2) = \mathsf{E}\{e^{t_1 X + t_2 Y}\} = \mathsf{E}\{\mathsf{E}[e^{t_1 X + t_2 Y} \mid X = x]\}$

$$= \mathsf{E}\{e^{t_1 X}\, \mathsf{E}(e^{t_2 Y} \mid x)\} = \mathsf{E}\{e^{t_1 X} . e^{Xt_2 + \frac{1}{2}\sigma^2 t_2^2}\}, \qquad [M(t : Y \mid x) = \text{m.g.f. of } N(x, \sigma^2)]$$

$$= e^{\frac{1}{2}\sigma^2 t_2^2}\, \mathsf{E}\{e^{(t_1 + t_2)X}\} = e^{\frac{1}{2}\sigma^2 t_2^2} . \exp\{\mu(t_1 + t_2) + \tfrac{1}{2}\sigma^2 (t_1 + t_2)^2\}$$

$$= e^{\mu(t_1 + t_2) + \frac{1}{2}\sigma^2 (t_1^2 + 2t_2^2 + 2t_1 t_2)}$$

This shows that $(X, Y) \sim$ BVN $(\mu, \mu\,; \sigma^2, 2\sigma^2;\ 1/\sqrt{2})$. $\qquad (\rho = 1/\sqrt{2})$

***Example* 4.** If $(X_i, Y_i) \sim$ BVN $(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2; \rho)$, $i = 1, 2, 3, \ldots, n$ is a random sample of size n, find the joint p.d.f. of $\overline{X} = (\Sigma X_i)/n$ and $\overline{Y} = (\Sigma Y_i)/n$.

Solution. Let $m(t_1, t_2)$ be the m.g.f. of $\overline{X}$ and $\overline{Y}$ and $M(t, t')$ that of (X_i, Y_i).

$$m(t_1, t_2) = \mathsf{E}[\exp(t_1 \overline{X} + t_2 \overline{Y})] = \mathsf{E}\{\exp[\Sigma(t_1 X_i / n) + (t_2 Y_i / n)]\}, \quad [t_1/n = \beta_1,\ t_2/n = \beta_2]$$

$$= \mathsf{E}\left\{\prod_{i=1}^{n} \exp(\beta_1 X_i + \beta_2 Y_i)\right\} = \Pi \mathsf{E}(e^{\beta_1 X_i + \beta_2 Y_i}) = \Pi M(\beta_1, \beta_2) = [M(\beta_1, \beta_2)]^n \{\because (X_i, Y_i) \text{ are i.i.d. vectors}\}$$

$$= \left[\exp(\mu_1 \beta_2 + \mu_2 \beta_2 + \tfrac{1}{2}(\sigma_1^2 \beta_1^2 + 2\rho\sigma_1\sigma_2\beta_1\beta_2 + \sigma_2^2\beta_2^2))\right]^n$$

$$= \exp\left\{\mu_1 t_1 + \mu_2 t_2 + \tfrac{1}{2}\left[\frac{\sigma_1^2}{n} t_1^2 + 2\rho \frac{\sigma_1 \sigma_2}{n} t_1 t_2 + \frac{\sigma_2^2}{n} t_2^2\right]\right\}$$

This shows that $(\overline{X}, \overline{Y}) \sim$ BVN $(\mu_1, \mu_2, \sigma_1^2/n, \sigma_2^2/n; \rho)$.

***Example* 5.** Let U, V, W be i.i.d. $N(0, 1)$ variates and T be $U(0, 1)$ variate. Define

$$X = U\sqrt{T} + V\sqrt{1 - T}, \qquad Y = W\sqrt{T} + V\sqrt{1 - T}.$$

Show that $X \sim N(0, 1)$, $Y \sim N(0, 1)$ but (X, Y) is not BVN.

Solution. $M(t : X) = \mathsf{E}(\exp tX) = \mathsf{E}[\exp(t\sqrt{T}U + t\sqrt{1 - T}V)]$

$$= \mathsf{E}_T\{\mathsf{E}(e^{(t\sqrt{T})U}) . \mathsf{E}(e^{t\sqrt{1-T}V}) \mid T = \text{Const.}\} \qquad [\text{by Double-E Rule}]$$

$$= \mathsf{E}[e^{t^2 T/2} . e^{t^2(1-T)/2}] = \mathsf{E}(e^{t^2/2}) = e^{t^2/2}. \qquad \ldots(1)$$

Thus, $X \sim N(0, 1)$. Similarly, $Y \sim N(0, 1)$.

$$M(t_1, t_2) = \mathsf{E}[\exp(t_1 X + t_2 Y)] = \mathsf{E}\{\exp[t_1 \sqrt{T}U)]\, \mathsf{E}[exp(t_2\sqrt{T}W)]\, \mathsf{E}[exp[(t_1 + t_2)\sqrt{1 - T}V)] \mid T = \text{Const.}\}$$

$$= \mathsf{E}_T\{\exp(\tfrac{1}{2}Tt_1^2 + \tfrac{1}{2}Tt_2^2 + \tfrac{1}{2}(1 - T)(t_1 + t_2)^2\} = \exp\{(t_1 + t_2)^2/2\}\, \mathsf{E}_T(e^{-t_1 t_2 T})$$

$$= \exp[(t_1 + t_2)^2/2]\ (1 - e^{-t_1 t_2})/t_1 t_2.$$

This is not the m.g.f. of BVN distribution hence (X, Y) is not jointly normal.

Remark. Put $t_2 = 0$, use L'Hospital Rule for Indeterminate to recover

$$M(t_1, 0) = \exp\left(\tfrac{1}{2}t^2\right) \Rightarrow X \sim N(0,1).$$

Evaluation of (1) is to emphasize an excellent technique.

***Example* 6.** Show that for standard BVN (0, 0, 1, 1 ; ρ) moment recurrence relation is

$$\mu_{r,s} = (r+s-1)\rho\mu_{r-1,s-1} + (r-1)(s-1)(1-\rho^2)\mu_{r-2,s-2}.$$

Hence or otherwise show that

$$\mu_{3,1} = 3\rho, \mu_{2,2} = 1+2\rho^2, \ \mu_{r,s} = 0, \text{ if } r+s \text{ is odd.}$$

Solution. $M \equiv M(t_1, t_2) = \exp\left\{\frac{1}{2}(t_1^2 + 2\rho t_1 t_2 + t_2^2)\right\} = \sum\limits_{r=0}^{\infty}\sum\limits_{s=0}^{\infty} \frac{\mu_{r,s}\, t_1^r t_2^s}{r\, s\,!}$. ...(1)

Let suffixes indicate partial differention, e.g. $M_1 = \partial M/\partial t_1$, $M_{12} = \partial^2 M/\partial t_2 \partial t_2$, etc.

$$\therefore \quad M_1 = (t_1 + \rho t_2)M, \quad M_2 = (t_2 + \rho t_1)M,$$

$$M_{12} = [\rho + (t_2 + t_1\rho)(t_1 + \rho t_2)]M = [\rho + (1+\rho^2)t_1 t_2 + \rho(t_1^2 + t_2^2)]M.$$

These relations immediately provide, on eliminating t_1^2, t_2^2

$$M_{12} = [\rho t_1 M_1 + \rho t_2 M_2 + \rho M] + (1-\rho^2)t_1 t_2 M. \qquad ...(2)$$

We now use series (1), differentiating and substituting into (2)

$$\sum_{r,s} \frac{\mu_{r,s} t_1^{r-1}\, t_2^{s-1}}{(r-1)!(s-1)!} = \rho \sum_{r,s} \frac{\mu_{r,s} t_1^r\, t_2^s}{r!s!}(r+s+1) + (1-\rho^2)\sum_{r,s} \frac{\mu_{r,s} t_1^{r+1}\, t_2^{s+1}}{r!s!}$$

Equating coeff. of $t_1^{r-1} t_2^{s-1}/(r-1)!(s-1)!$ in the above Eq. we get

$$\mu_{r,s} = \rho\{(r-1)+(s-1)+1\}\mu_{r-1,s-1} + (1-\rho^2)(r-1)(s-1)\mu_{r-2,s-2}$$

$$= (r+s-1)\rho\mu_{r-1,s-1} + (r-1)(s-1)(1-\rho^2)\mu_{r-2,s-2}. \qquad ...(3)$$

(ii) Since X and Y are standard, $\sigma_1 = \sigma_2 = 1$. Also $\mu_{1,1} = \rho\sigma_1\sigma_2 = \rho$; $\mu_{0,0} = 1$, so

$$\mu_{3,1} = 3\rho\mu_{2,0} - 3\rho, \ \mu_{2,2} = 3\rho\mu_{1,1} + (1-\rho^2)\mu_{0,0} = 3\rho^2 + (1-\rho^2) = 1+2\rho^2.$$

To prove that $\mu_{r,s} = 0$, if $r + s$ is odd, we notice that

$$\mu_{3,0} = \mu_{0,3} = 0, \ \mu_{1,2} = \mu_{2,1} = 0, \text{ as } \mu_{0,1} = \mu_{1,0} = 0.$$

So we proceed by Mathematical induction.

If $r + s$ is odd, then $(r-1) + (s-1)$, $(r-2) + (s-2)$, $(r-3) + (s-3)$ etc. are all odd. Assume that $\mu_{r-1,s-1} = 0$, $\mu_{r-2,s-2} = 0$. Now by (3), $\mu_{r,s} = 0$; hence the inductive process is complete and the result fully established.

***Example* 7.** If $f(x, y)$ is the p.d.f. of BVN $(0, 0, 1, 1; \rho)$, show that $\partial f/\partial \rho = \partial^2 f/\partial x\,\partial y$. If $V = P(0 < T \le x)$, $W = P(0 < T \le y)$, where T is $N(0, 1)$; prove that W and V are uniform over $(-1/2, 1/2)$. Prove further that $R = \text{Corr}\,(W, V)$ satisfies the relation

$$\rho = 2 \sin (\pi R/6).$$

Solution. Here $\quad f(x, y) = (2\pi\sqrt{1-\rho^2})^{-1} \exp[-(x^2 - 2\rho xy + y^2)/2(1-\rho^2)], \qquad \ldots(1)$

$$\therefore \quad \ln f(x, y) = -\ln(2\pi\sqrt{1-\rho^2}) - (x^2 - 2\rho xy + y^2)/2(1-\rho^2).$$

$$\frac{\partial f}{\partial x} = -\frac{(x-\rho y)}{1-\rho^2} f ; \frac{\partial f}{\partial y} = -\frac{(y-\rho x)}{1-\rho^2} f ; \quad \frac{\partial^2 f}{\partial x\,\partial y} = \left[\frac{\rho}{1-\rho^2} + \frac{(x-\rho y)(y-\rho x)}{(1-\rho^2)^2}\right] f = \frac{\partial f}{\partial \rho}$$

Recall : $F(X) \sim U(0, 1)$ and $Z = [(b-a) F(X) + a] \sim U(a, b)$.

Here $\quad F_T(x) = (1/2) + V \Rightarrow V = [(1/2) - (-1/2) F + (-1/2) \quad \Rightarrow V \sim U(-1/2, 1/2)$.

Now $\qquad E(V) = E(W) = [1/2 + (-1/2)]/2 = 0$, Var $V = (b-a)^2/12 = 1/12$.

$$R = \sigma_{VW}/\sigma_V.\sigma_W = 12\,E(VW) = 12 \iint wv\, f(x, y)\, dx\, dy, \quad [w \,\&\, v \text{ are dep. function of } x, y]$$

$$\frac{dR}{d\rho} = 12 \iint wv \frac{\partial f}{\partial \rho} dx\, dy = 12 \iint wv \frac{\partial^2 f}{\partial x\,\partial y} dx\, dy.$$

We integrate by parts w.r.t. x and y and obtain

$$\frac{dR}{d\rho} = 12 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{dw}{dx} \frac{dv}{dy} f(x, y)\, dx\, dy, \quad \left[w = \int_0^x \frac{e^{-t^2/2}}{\sqrt{2\pi}} dt, \ \frac{dw}{dx} = \frac{e^{-x^2/2}}{\sqrt{2\pi}}\right]$$

$$= \frac{3}{\pi^2\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp\{-[(2-\rho^2)(x^2+y^2) - 2\rho xy]/2(1-\rho^2)\}\, dx\, dy, \qquad [\text{by } (1)].$$

Recall : $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp[-(ax^2 + 2hxy + by^2)]\, dx\, dy = \dfrac{\pi}{\sqrt{ab-h^2}}$

$$\therefore \quad \frac{dR}{d\rho} = \frac{3}{\pi^2\sqrt{(1-\rho^2)}} \frac{\pi\, 2(1-\rho^2)}{[(2-\rho^2)^2 - \rho^2]^{1/2}} = \frac{6\sqrt{1-\rho^2}}{\pi\sqrt{(1-\rho^2)(4-\rho^2)}} = \frac{6}{\pi\sqrt{4-\rho^2}}.$$

Integration gives : $R = (6/\pi) \sin^{-1} (\rho/2) + C = (6/\pi) \sin^{-1} (\rho/2)$,
since when $\rho = 0$, $R = 0$, hence $C = 0$. Thus $\rho = 2 \sin (\pi R/6)$.

***Example* 8.** Let W and Z be two random numbers chosen from the unit interval. Show that

$$X = \sqrt{-2\ln W}\left[\frac{\cos 2\pi Z}{\sqrt{2(1-\rho)}} + \frac{\sin 2\pi Z}{\sqrt{2(1+\rho)}}\right], \ Y = \sqrt{-2\ln W}\left[\frac{\cos 2\pi Z}{\sqrt{2(1-\rho)}} - \frac{\sin 2\pi Z}{\sqrt{2(1+\rho)}}\right]$$

are dependent $N(0, 1)$ variates with correlation coefficient ρ.

Solution. For neatness, let $a = [2(1+\rho)]^{-1/2}$, $\quad b = [2(1-\rho)]^{-1/2}$, $\quad 2\pi z = \alpha$, so that

$$x = \sqrt{-2\ln w}\,(b\cos\theta + a\sin\theta) ; \ y = \sqrt{-2\ln w}\,(b\cos\theta - a\sin\theta)$$

$$J = \frac{\partial(x, y)}{\partial(w, z)} = \frac{\sqrt{-2\ln w}\,(2\pi)\sqrt{-2}}{2w\sqrt{\ln w}} \begin{vmatrix} b\cos\theta + a\sin\theta & b\cos\theta + a\sin\theta \\ a\cos\theta - b\sin\theta & -b\sin\theta - a\cos\theta \end{vmatrix}$$

Add column 2 to column 1 to get

$$\frac{\partial(x,y)}{\partial(w,z)} = \frac{4\pi b}{w}\begin{vmatrix}\cos\theta & b\cos\theta - a\sin\theta\\ \sin\theta & a\cos\theta + b\sin\theta\end{vmatrix} = \frac{4\pi ab}{w}\begin{vmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{vmatrix} = \frac{4\pi ab}{w}.$$

This provides $dx\,dy = (4\pi\, ab/w)\,dz\,dw$, hence the joint p.d.f. of variates $[f(w, z) = 1]$ is

$$1. \qquad dw\,dz = (w\sqrt{1-\rho^2}/2\pi)\,dx\,dy. \qquad \ldots(1)$$

Now, $x^2 + y^2 = (-4\ln w)(a^2\sin^2\theta + b^2\cos^2\theta) = -2(\ln w)(1+\rho\cos 2\theta)/(1-\rho^2)$. ...(2)

$$xy = (-2\ln w)(b^2\cos^2\theta - a^2\sin^2\theta) = -(\ln w)(\rho + \cos 2\theta)/(1-\rho^2). \qquad \ldots(3)$$

Also $\quad \mathrm{E}[\cos(2\pi Z)] = \int_0^1 \cos 2\pi\,dz = 0 = \mathrm{E}[\sin(2\pi Z)]$ ...(4)

$$\mathrm{E}[\cos^2(2\pi Z)] = \left(\frac{1}{2}\right)\int_0^1 (1 + 4\cos 4\pi z)\,dz = 1/2 = \mathrm{E}[\sin^2(2\pi Z)] \qquad \ldots(5)$$

From independence of W and Z, using (4), we conclude that $\mathrm{E}(X) = 0 = \mathrm{E}(Y)$.

Also $\quad \int_0^1 \ln w\,dw = [w\ln w - w]_0^1 = -1.$ ...(6)

$\therefore \quad \mathrm{E}(XY) = -(1-\rho^2)^{-1}\,\mathrm{E}(\ln W).\mathrm{E}(\rho + \cos 4\pi Z) = \rho/(1-\rho^2)$ [Using (3)] ...(7)

$$\sigma_{XY} = \mathrm{E}(XY) - \mathrm{E}(X)\mathrm{E}(Y) = \mathrm{E}(XY) \neq 0 \Rightarrow X \text{ and } Y \text{ are dependent.}$$

Now, $\quad X^2 = (-\ln W)(b^2\cos^2\theta) + a^2\sin^2\theta + ab\sin 2\theta)$

$\therefore \quad \mathrm{E}(X^2) = -2\mathrm{E}(\ln W)\,\mathrm{E}[b^2\cos^2\theta + a^2\sin^2\theta + ab\sin 2\theta]$ [Use (5) and (6)]

$$= (-2)(-1)[\tfrac{1}{2}b^2 + \tfrac{1}{2}a^2 + 0] = a^2 + b^2 = 1/(1-\rho^2).$$

$\therefore \quad \mathrm{Var}(X) = \mathrm{Var}(Y) = 1/(1-\rho^2);\ \mathrm{Corr}(X,Y) = \sigma_{XY}/\sigma_X\sigma_Y = \rho.$ [by (7)]. ...(8)

Now : $\quad \dfrac{x^2}{\sigma_X^2} + \dfrac{y^2}{\sigma_Y^2} - \dfrac{2xy}{\sigma_X\sigma_Y}\rho = (1-\rho^2)(x^2 + y^2 - 2\rho xy)$

$\therefore Q(x,y) = (x'^2 + y'^2 - 2\rho xy)/(1-\rho^2) = z^2 + y^2 - 2\rho x'y' = -2\ln w\,[x' = x/\sigma_1, \text{etc.}]$ [by (2) & (3)]

This gives $w = e^{-\frac{1}{2}Q(x,y)}$ and substituting in (1) we get

$$f(x,y) = \frac{1}{2\pi}\sqrt{1-\rho^2}\,e^{-Q(x,y)/2} = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\,e^{-Q(x,y)/2} \qquad \ldots(9)$$

This proves that X and Y are jointly normal with correlation ρ.

Problems with Solutions Provided at the End of the Text

1*. Let $(X, Y) \sim \mathrm{BVN}(0, 0, \sigma_1^2, \sigma_2^2; \rho)$. Show that $U = (X/\sigma_1) - (\rho Y/\sigma_2)$, $V = Y\sqrt{1-\rho^2}/\sigma_2$ are independent Normal variates.

2*. Determine the parameters for the BVN distribution

$$f(x,y) = C\exp\{-[16(x-2)^2 - 12(x-2)(y+3) + 9(y+3)^2]/216\}.$$

3*. Find Corr (X, Y) for jointly Normal distribution

$$f(x, y) = (2\pi\sqrt{3})^{-1} \exp\{-[(2x - y)^2 + 2xy)/6]; \quad -\infty < x, y < \infty \qquad (1)$$

4*. Show that if (X, Y) is BVN $(0, 0, 1, 1\ ;\ \rho)$, then

$$P\{XY > 0\} = \tfrac{1}{2} + \pi^{-1}\sin^{-1}(\rho), \qquad P\{XY < 0\} = \pi^{-1}\cos^{-1}(\rho)$$

5*. If (X, Y) is BVN $(5, 10, 1, 25;\ \rho)$, find $\rho(> 0)$ when $P\{4 < Y < 16 \mid X = 5\} = 0.954$
If $\rho = 0$, find $P\{X + Y \leq 16\}$.

6*. If (X, Y) is BVN $(3, 1, 16, 25;\ 3/5)$, find

(a) $P(3 < Y < 8 \mid x = 7)$, (b) $P(-3 < X < 3 \mid y = -4)$.

7*. Random variables X and Y have joint p.d.f.

$$f(x, y) = (6\pi\sqrt{3})^{-1} \exp\left\{-\frac{2}{3}\left[\frac{1}{9}x^2 - \frac{5}{6}x + \frac{1}{6}xy + \frac{1}{4}y^2 - y + \frac{7}{8}\right]\right\}; \quad -\infty < x, y < \infty.$$

Find the best M.S. estimate of Y using X and the M.S. error involved with using this estimate.

8*. If (X, Y) is BVN $(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2\ ;\ \rho)$, compute Corr (e^X, e^Y).

9*. Let X, Y, Z be i.i.d. $N(0, 1)$ variates and suppose $U = X + \lambda Z$, $V = Y + \lambda Z$

(i) Show that U and V have BVN distribution and find λ so that the Corr Coeff. is 1/2.

(ii) With ρ being the same as before, what additional transformation involving U and V would produce a BVN $(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2;\ \rho)$?

10*. Let both X and Y be mixed variates such that

$$f(x, y) = k(2\pi\sigma_1 . \sigma_2\sqrt{1-\rho^2})^{-1} \exp[-Q/2(1-\rho^2)]; \quad -\infty < x, y < \infty, \text{ integers omitted}$$

where $Q = [(x-\mu_1)^2/\sigma_1^2] - 2\rho[(x-\mu_1)(y-\mu_2)/\sigma_1\sigma_2] + [(y-\mu_2)^2/\sigma_2^2]$; $f(x, y) = 0$, otherwise;

$$p(x, y) = \binom{n}{x}\binom{n-x}{y} p^x q^y (1-p-q)^{n-x-y}(1-k); \quad x, y \text{ integers}$$

Find the m.g.f. for the pair (X, Y), Var (X), Var $(X \mid y)$ and Cov (X, Y).

11*. Let : $f(x, y) = (2\pi)^{-1} e^{-(x^2+y^2)/2}[1 + xye^{-(x^2+y^2-2)/2}], \quad -\infty < x,\ y < \infty.$

Verify that $f(x, y)$ is a p.d.f. and that each marginal p.d.f. is Normal. Are X and Y independent ?

12*. The distribution of two components measurement (X, Y) is BVN $(\mu_1, \mu_2, 1, 1\ ;\ \rho)$. For the proper functioning of a machine, which uses the components, it is necessary that some linear combination of X and Y can be obtained with unit variance and independent of X. Is such a combination possible ?

13*. If $X \sim N(0, \sigma_1^2)$ and $Y \sim N(0, \sigma_2^2)$ have correlation coefficient ρ, show that Corr $(X^2, Y^2) = \rho^2$.

14*. For $(X, Y) \sim$ BVN $(0, 0, 1, 1, \rho)$, show that

(a) $(X + Y)$ and $(X - Y)$ are independent Gaussian distributed.

(b) X and Y are independent iff $\rho = 0$. (c) $Q = (X^2 - 2\rho XY + Y^2)/\sqrt{1-\rho^2}$ has $\chi^2_{(2)}$ distribution.

15*. Variates X and Y with zero means and S.D.s σ_1, σ_2 are normally correlated with correlation coefficient ρ. Show that

$$U = (X/\sigma_1) + (Y/\sigma_2), \; V = (X/\sigma_1) - (Y/\sigma_2)$$

are independent $N[0, 2(1+\rho)]$ and $N[0, 2(1-\rho)]$ variates.

16*. Let $(X, Y) \sim$ BVN $(0, 0, \sigma_1^2, \sigma_2^2; \rho)$. Show that, using conditional distributions,

$$\text{Var}(XY) = \sigma_1^2\sigma_2^2(1+\rho^2), \qquad \text{Corr}(X^2, Y^2) = \rho^2.$$

17*. Let $Q(\rho) = (x^2 - 2\rho xy + y^2)/(1-\rho^2)$, $-\infty < x, \; y < \infty$ and define

$$f(x, y) = (4\pi\sqrt{1-\rho^2})^{-1}\{\exp[-\tfrac{1}{2}Q(\rho)] + \exp[-\tfrac{1}{2}Q(-\rho)]\}. \qquad ...(1)$$

Show that $f(x, y)$ is a joint p.d.f. such that both marginals are $N(0, 1)$ but $f(x, y)$ is not BVN. Show further that X and Y have zero correlation although X and Y are not independent.

18*. Show that the normality of conditional densities does not imply that the bivariate density is normal (Gaussian).

19*. If $(X, Y) \sim$ BVN $(0, 0, 1, 1; \rho)$, find E [max (X, Y)] and E[min (X, Y)].

20-70. An Alternate Development of BVN Distribution

Def. A random vector (X_1, X_2) is said to possess BVN distribution, if for *all* a and b, $Z = aX_1 + bX_2$, is a normal variate.

1. *Moment Generating Function.* Let $E(X_i) = \mu_i$, Var $(X_i) = \sigma_i^2$, $i = 1, 2$. Then

$$\mu = E(Z) = a\mu_1 + b\mu_2, \sigma^2 = \text{Var}(Z) = a^2\sigma_1^2 + b^2\sigma_2^2 + 2ab\,\rho\sigma_1\sigma_2.$$

Since $Z \sim N(\mu, \sigma^2)$, $M(t : Z) = \exp[\mu t + \tfrac{1}{2}\sigma^2 t^2]$, i.e.

$$M(t : Z) = \exp\{(a\mu_1 + b\mu_2)t + \tfrac{1}{2}(a^2\sigma_1^2 + 2ab\rho\sigma_1\sigma_2 + b^2\sigma_2^2)t^2\}.$$

or $$M(t_1, t_2) = \exp\{\mu_1 t_1 + \mu_2 t_2 + \tfrac{1}{2}(\sigma_1^2 t_1^2 + 2\rho\sigma_1\sigma_2 t_1 t_2 + \sigma_2^2 t_2^2)\} \qquad ...(1)$$

where $at = t_1$ and $bt = t_2$. Result (1) is the m.g.f. of BVN-distribution.

2. *Independence versus Correlation.* This is proved in § 20-50, no change in new setup.

3. *Probability density function for BVN distribution.* Let $\rho^2 < 1$ and define :

$$U = (X_1 - \mu_1)/\sigma_1, \; V = (X_2 - \mu_2)/\sigma_2, \; W = (V - \rho U)/\sqrt{1-\rho^2}.$$

Obviously, $E(U) = E(V) = E(W) = 0$, Var (U) = Var (V) = Var $(W) = 1$, as $\rho = E(UV)$, Cov $(U, W) = 0$.

Since U and V are linear combinations of X_1 and X_2, they have a BVN distribution. Also Cov $(U, W) = 0 \Rightarrow U$ and W are independent [characteristics property of BVN-distribution] $N(0, 1)$ variates, and hence their joint elemental p.d.f. is

$$dF(u, w) = (2\pi)^{-1} \exp[-\tfrac{1}{2}u^2 - \tfrac{1}{2}w^2]\,du\,dw.$$

Since $$\frac{\partial(u, w)}{\partial(x_1, x_2)} = \begin{vmatrix} \sigma_1^{-1} & 0 \\ -\rho(\sigma_1\sqrt{[1-\rho^2]^{-1}} & [\sigma_2\sqrt{[(1-\rho^2)]^{-1}} \end{vmatrix} = \frac{1}{\sigma_1\sigma_2\sqrt{1-\rho^2}}$$

$$\therefore \quad dF(x_1, x_2) = (2\pi\sigma_1\sigma_2\sqrt{1-\rho^2})^{-1} \exp[-\tfrac{1}{2}Q]\,dx_1\,dx_2 \quad \ldots(1)$$

where $Q = u^2 + w^2 = (u^2 - 2\rho uv + v^2)/(1-\rho^2)$

$$= \frac{1}{1-\rho^2}\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_2-\mu_2}{\sigma_2}\right) + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right]. \quad \ldots(2)$$

From (1) & (2) we get the required p.d.f. for BVN $(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2; \rho)$:

$$f(x_1, x_2) = (2\pi\sigma_1\sigma_2\sqrt{1-\rho^2})^{-1} \exp[-\tfrac{1}{2}Q], \ -\infty < x_1, x_2 < \infty. \quad \ldots(3)$$

Note. Let $Y_1 = aX_1 + bX_2, Y_2 = cX_1 + dX_2$; then $\alpha Y_1 + \beta Y_2$ is a linear combination of X_1 and X_2 and hence (Y_1, Y_2) has BVN distribution.

Example : Let X and Y be standardized variates. Let $Z = aX + bY$, $(a \neq 0, b \neq 0)$ be $N(0, a^2 + ab + b^2)$. Obtain the joint distribution of (X, Y).

Solution. A characterization property of normal distribution states that if $Z = aX + bY + c$ is Normal, then X & Y are jointly Normal. Here

$$\text{Var } Z = a^2\sigma_X^2 + b^2\sigma_Y^2 + 2ab\,\sigma_X\sigma_Y\rho_{XY} = a^2 + b^2 + 2ab\rho.$$

Thus by hypothesis, $2\rho = 1$. Hence the joint density of (X, Y) is BVN (0, 0, 1, 1, ; 1/2).

i.e. $$f(x, y) = (\pi\sqrt{3})^{-1} \exp[-2(x^2 + y^2 - xy)/3], \ -\infty < x, y < \infty.$$

20-71. *Linear Functions of Normal Variates*

Theorem. If normal variates X and Y have the joint p.d.f.

$$f(x, y) = (2\pi\sqrt{1-\rho^2})^{-1} \exp[-(x^2 - 2\rho xy + y^2)/2(1-\rho^2)]; \ -\infty < x, y < \infty$$

then any linear function $U = aX + bY$ of these two variates is also Normally distributed.

Proof. We apply transformations of the type : $u = ax + by$, $v = a'x + b'y$ so chosen that $K(u^2 + v^2) = x^2 - 2\rho xy + y^2, (K = \text{const.})$

$$\therefore \quad P\{U \leq u\} = \iint_{u<v} f(x, y)\,dx\,dy = \int_{-\infty}^{u}\int_{-\infty}^{\infty} e^{-K(u^2+v^2)}\,du\,dv = B\int_{-\infty}^{u} e^{-Ku^2}\,du.$$

If follows that U is normally distributed.

Since $E(U) = 0$; Var $(U) = a^2 + b^2 + 2ab\rho$, hence the p.d.f. of U is

$$f_1(u) = [2\pi(a^2 + b^2 + 2ab\rho)]^{-1/2} \exp[-u^2/2(a^2 + b^2 + 2ab\rho)]; \ -\infty < u < \infty.$$

Note. The deviation of the extension to general value of means and variances is straight forward. [Apply m.g.f. method also].

Cor. Any linear function of two independent Normal variates is normally distributed.

Remark. Two normal variates may have joint distribution which is *not* BVN, in this case a linear function of these variables need not be normal [Vide Example 20-19(a)].

20-72. Cauchy Theorem : Ratio of Two Correlated Normal Variates

Let $(X, Y) \sim$ BVN $(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2 ; \rho)$, then the distribution of

$$Z = \frac{(X-\mu_1)/\sigma_1}{(Y-\mu_2)/\sigma_2} = \frac{U}{V}$$

is $f(z) = \sqrt{1-\rho^2}/\pi[z^2 - 2\rho x + 1], \; -\infty < z < \infty.$

Proof by Double-E Rule. Let $U = (X-\mu_1)/\sigma_1$, $V = (Y-\mu_2)/\sigma_2$, $W = (V-\rho U)/\sqrt{1-\rho^2}$.

Then U and W are indep. $N(0, 1)$ variates. Let $Z = V/U$ or $Z = U/V$ is similar. We offer *two* evaluations.

(i)

$$\begin{aligned}\phi(t:Z) &= \mathsf{E}(e^{itZ}) = \mathsf{E}(e^{itV/U}) = \mathsf{E}\{\mathsf{E}\,(e^{itV/U}\; U=u) = \mathsf{E}\{\mathsf{E}e^{(it/u)V} \mid u=u\} \\ &= \mathsf{E}\{e^{i(\rho u)(t/u) - \frac{1}{2}(1-\rho^2)t^2/u^2} \mid u\},\end{aligned}$$

$[Y \mid x \sim N[\mu_2 + \rho(\sigma_2/\sigma_1)(X-\mu_1); \sigma_2^2(1-\rho^2)]$; for $V \mid u$, use $\mu_2 = 0 = \mu_1, \sigma_1 = \sigma_2 = 1$

$$\begin{aligned}\therefore\quad \phi(t:Z) &= e^{i\rho t}\,\mathsf{E}\{e^{-(1-\rho^2)t^2/2u^2}\} \\ &= e^{i\rho t}\int_{-\infty}^{\infty} \frac{e^{-\frac{1}{2}u^2}}{\sqrt{2\pi}}\, e^{-(1-\rho^2)t^2/2u^2}\, du, \qquad \left[\int_0^{\infty} e^{-a^2x^2 - (b^2/x^2)}\, dx = \frac{\sqrt{\pi}}{2a}\, e^{-2ab}, a > 0, b \geq 0\right] \\ &= \frac{2e^{i\rho t}}{\sqrt{2\pi}}\int_0^{\infty} e^{-(u^2/2) - [(1-\rho)^2 t^2/2u^2]}\, du. \qquad \left[a^2 = 1/2,\; b = \sqrt{1-\rho^2}\,|t|/2\right] \\ &= e^{i\rho t - |t|\sqrt{1-\rho^2}}.\end{aligned}$$

(ii)

$$\begin{aligned}\phi(t:Z) &= \mathsf{E}(e^{itZ}) = \mathsf{E}(e^{itV/U}) = \mathsf{E}[e^{it(\rho U + \sqrt{1+\rho^2}W)/U}] = e^{it\rho}\,\mathsf{E}\{e^{it\sqrt{1-\rho^2}W/U}\} \\ &= e^{it\rho}\,.\mathsf{E}\{e^{it\sqrt{1-\rho^2}W/U} \mid U=u\} = e^{it\rho}\,\mathsf{E}\{e^{-\frac{1}{2}t^2(1-\rho^2)/U^2}\}\;[\phi_w(t) = e^{-t^2/2}] \\ &= e^{i\rho t}\int_{-\infty}^{\infty} \frac{e^{-\frac{1}{2}u^2}}{\sqrt{2\pi}}\, e^{-t^2(1-\rho^2)/2u^2}\, du = \frac{2e^{i\rho t}}{\sqrt{2\pi}}\int_0^{\infty} e^{-(u^2/2) - [t^2(1-\rho^2)/2u^2]}\, du \\ &= e^{i\rho t - |t|\sqrt{1-\rho^2}}. \qquad \text{[as above]}\end{aligned}$$

This is the Ch. Function of a Cauchy variate. Since Ch. Function uniquely identifies a distribution, it follows that

$$f(z) = \sqrt{1-\rho^2}/\pi[(z-\rho)^2 + (\sqrt{1-\rho^2})^2], \quad -\infty < z < \infty.$$

Comments. Evaluation (ii) is neater, it avoids the conditional distribution used in (i).

Note. When $\rho = 0$, this theorem corresponds to theorem 2, § 19-58.

Cor. Distribution of $T = (X - \mu_1)/(Y - \mu_2)$.

Put $t = \sigma_1 z/\sigma_2$, i.e. $z = t(\sigma_2/\sigma_1)$; $dz = (\sigma_2/\sigma_1)\,dt$ and (1) reduces to

$$g(t) = \frac{\sqrt{1-\rho^2}\,(\sigma_2/\sigma_1)}{\pi\{(\sigma_2/\sigma_1)^2 t^2 - 2\rho(\sigma_2/\sigma_1)t + 1\}} = \frac{\sigma_1\sigma_2\sqrt{1-\rho^2}}{\pi[\sigma_2^2 t^2 - 2\rho\sigma_1\sigma_2 t + \sigma_1^2]}, \quad -\infty < t < \infty.$$

$$= \lambda/\pi\{[t - \rho(\sigma_1/\sigma_2)]^2 + \lambda^2\}, \quad \lambda = \sigma_1\sqrt{1-\rho^2}/\sigma_2, \quad -\infty < t < \infty.$$

This is Chy $(\rho\sigma_1/\sigma_2, \lambda)$, with modal value $\rho(\sigma_1/\sigma_2)$.

20-80. *Price Theorem*

If (X, Y) is BVN $(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2; \rho)$, then for any $g(X, Y)$: an arbitrary function of X and Y, and $c = \sigma_{12} = \sigma_1\sigma_2\rho$.

$$\frac{\partial^n}{\partial c^n}\mathsf{E}\{g(X,Y)\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{\partial^{2n} g(x,y)}{\partial x^n\,\partial y^n} f(x,y)\,dx\,dy = \mathsf{E}\left\{\frac{\partial^{2n} g(X,Y)}{\partial X^n\,\partial Y^n}\right\}.$$

Proof. By Inversion Theorem, p.d.f. in terms of Ch. Function is given by

$$f(x,y)\frac{1}{4\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-i(t_1x+t_2y)}\,\phi(t_1,t_2)\,dt_1\,dt_2 \qquad ...(1)$$

where $$\phi(t_1,t_2) = \exp\{i(t_1\mu_1 + t_2\mu_2) - \tfrac{1}{2}(\sigma_1^2 t_1^2 + 2\rho\,\sigma_1\sigma_2\,t_1 t_2 + \sigma_2^2 t_2^2)\} \qquad ...(2)$$

$$\therefore \quad \mathsf{E}[g(X,Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)\,.\,\frac{1}{4\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-i(t_1+t_2x)}\,\phi(t_1,t_2)\,dt_1\,dt_2\,dx\,dy.$$

Since only $\phi(t_1,t_2)$ depends on $c = \sigma_{12} = \sigma_1\sigma_2\rho$; differentiating the above equation w.r.t. "c" n times and using $(\partial^n\phi/\partial c^n) = (-t_1 t_2)^n\phi$, we get

$$\frac{\partial^n \mathsf{E}[g(X,Y)]}{\partial c^n} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{g(x,y)}{4\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-i(t_1x+t_2y)}\,.(-t_1,t_2)^n\,\phi(t_1,t_2)\,dt_1\,dt_2\,dx\,dy \qquad ...(3)$$

Also differentiating (1), n times w.r.t. 'x' and n times w.r.t. 'y' we get

$$\frac{\partial^{2n} f(x,y)}{\partial x^n\,\partial y^n} = \frac{1}{4\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(-t_1\,t_2)^n\,e^{-i(t_1x+t_2y)}\,\phi(t_1,t_2)\,dt_1\,dt_2 \qquad ...(4)$$

$$\therefore \quad \frac{\partial^n}{\partial c^n}\mathsf{E}\{g(X,Y)\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)\frac{\partial^{2n} f(x,y)}{\partial x^n\,\partial y^n}\,dx\,dy, \quad \text{[by (3) \& (4)]} \qquad ...(5)$$

We assume that for $|x| \to \infty$ or $|y| \to \infty$, the function $g(x,y) \to \infty$ too rapidly, i.e. if $|g(x,y)| < A(e^{x^n+y^n})$, $n < 2$, then integrating by parts repeatedly we get

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)\frac{\partial^{2n} f(x,y)}{\partial x^n\,\partial y^n}\,dx\,dy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\frac{\partial^{2n} g(x,y)}{\partial x^n\,\partial y^n}\,dx\,dy \qquad ...(6)$$

The stated result follows from Eqs. (5) and (6).

Example : Let $\sin\alpha = \rho$, $-\frac{1}{2}\pi < \alpha \le \frac{1}{2}\pi$. If (X, Y) is BVN $(0, 0, \sigma_1^2, \sigma_2^2; \rho)$, then $E(|XY|) = (2\sigma_1\sigma_2/\pi)(\cos\alpha + \alpha\sin\alpha)$.

Solution. In Price Theorem, put $g(x, y) = |xy|$ to get

$$\frac{\partial}{\partial c} E(|XY|) = E\left\{\frac{d}{dx}|x|\frac{d}{dy}|y|\right\} \qquad \ldots(1)$$

But
$$\frac{d}{dx}|x| = \begin{cases}1, & x>0\\ -1, & x<0\end{cases}, \quad \frac{d}{dy}|y| = \begin{cases}1, & y>0\\ -1, & y<0\end{cases}$$

$$\therefore \quad E\left\{\frac{d}{dx}|x|.\frac{d}{dy}|y|\right\} = 1.P(XY>0) - 1.P(XY<0). \qquad \ldots(2)$$

Since $P(XY > 0) = \frac{1}{2} + (\alpha/\pi)$, $P(XY < 0) = \frac{1}{2} - (\alpha/\pi)$, by Example 20-4 Eqs. (1) and (2) provide

$$\frac{\partial}{\partial c} E(|XY|) = \frac{2\alpha}{\pi} = \frac{2}{\pi}\sin^{-1}\left(\frac{c}{\sigma_1\sigma_2}\right). \qquad \left(\because \rho = \frac{c}{\sigma_1\sigma_2}\right)$$

Integrating this result we get

$$E(|XY|) = \frac{2}{\pi}\int_0^c \sin^{-1}\left(\frac{c}{\sigma_1\sigma_2}\right)dc + \{E(|XY|)\}_{c=0}. \qquad \ldots(3)$$

When $c = 0$, (X, Y are independent), so

$$E|XY| = E|X|.E|Y| = (\sqrt{2/\pi})\sigma_1(\sqrt{2/\pi})\sigma_2 = 2\sigma_1\sigma_2/\pi.$$

$$\int_0^c \sin^{-1}\left(\frac{c}{\sigma_1\sigma_2}\right) = \left[c\sin^{-1}\left(\frac{c}{\sigma_1\sigma_2}\right) + \sqrt{\sigma_1^2\sigma_2^2 - c}\right]_0^c = \sigma_1\sigma_2[\alpha\sin\alpha + \cos\alpha - 1]$$

Making substitutions into (3), the given result follows.

Exercises

1. If X and Y have joint p.d.f. of the form

   $f(x, y) = K\exp\{-\frac{1}{2}[a_{11}(x-b)^2 + 2a_{12}(x-b)(y-c) + a_{22}(y-c)^2]\}$, $-\infty < x, y < \infty$.

   Find K and also Corr (X, Y).

2. Find the *Five* parameters of the BVN density

   (i) $f(x, y) = (2\pi)^{-1}\exp[-\frac{1}{2}(2x^2 + y^2 + 2xy - 22x - 14y + 65]$

   (ii) $f(x, y) = (2\pi)^{-1}\exp[-\frac{1}{2}x^2 - \frac{1}{2}y^2 - 2x + 3y - (13/2)]$.

3. For the BVN $(1, 2, 4^2, 5^2; 12/13)$, find $P(X > 2)$ and $P(X > 2 \mid Y = 2)$.

4. For the BVN $(60, 75, 6^2, 12^2; 0.55)$; find
   (i) $P(65 < X < 75)$, (ii) $P(71 \le Y \le 80 \mid X = 55)$, (iii) $P(|X - Y| \ge 15)$.

5. For the BVN : $f(x, y) = k \exp [-(2/3)(x^2 - xy + y^2 - 3x + 3y + 3)]$, find the marginal distribution of Y and the conditional distribution of Y given x.
6. If (X, Y) is BVN $(0, 0, 1, 1 ; 0.8)$, find the p.d.f. of $Z = X - Y$ and identify it.
7. (a) If (X, Y) is BVN $(0, 0, 1, 1 ; \rho)$, prove that $P\{X > 0, Y > 0\} = \frac{1}{4} + (2\pi)^{-1} \sin^{-1}(\rho)$

   (b) Let $(X, Y) \sim$ BVN $(0, 0, 1, 1 ; 0)$. Determine the radius of the circle S with centre at $(0, 0)$ such that $P\{(X, Y) \in S\} = 0.95$.
8. (a) If (X, Y) is BVN $(0, 0, \sigma_1^2, \sigma_2^2 ; \rho)$, $(\lambda = \sigma_2 / \sigma_1)$, show that

$$p = P\{|X| \le |Y|\} = (\pi)^{-1} \tan^{-1} [2\lambda \sqrt{1 - \rho^2} / (1 - \lambda^2)].$$

   Hence show that, irrespective of the value of ρ in its permissible range of variation, $p >$ or $<$ $\frac{1}{2}$ according as $\sigma_1 <$ or $> \sigma_2$. Try to generalize the result.

   (b) Let $(X, Y) \sim$ BVN $(0, 0, \sigma^2, \sigma^2 ; \rho)$. Find the distribution of $Z = X/Y$ and use this result to show that $P\{X < 0, Y > 0\} = (1/2)\, P(XY < 0\} = (2\pi)^{-1} \cos^{-1} \rho$.
9. Let $f(x, y) = K \exp [-4x^2 - 6xy - 9y^2]$, $-\infty < x,\ y < \infty$.
   Find the constant K, Cov (X, Y) and the conditional distributions $f(x \mid y)$ and $f(y \mid x)$.
10. A manufacturer of quality electric bulbs rejects bulbs for which a certain quality characteristic X of flament is less than 65 units. The quality characteristic X and the life (in hours) Y obey BVN $(80, 1100, 10^2, 100^2, 0.6)$. Find

    (i) The proportion of bulbs produced that will burn for less than 1000 hours.

    (ii) The proportion of bulbs that can be put on sale.

    (iii) The average life of bulbs put for sale.
11. Let $f(x, y, z) = (2\pi)^{-3/2} \exp [-\frac{1}{2}(x^2 + y^2 + z^2)] \{1 + xyz \exp [-\frac{1}{2}(x^2 + y^2 + z^2)]\}$ where $-\infty < x, y, z < \infty$. While X, Y, Z are obviously stochastically dependent show that X, Y and Z are pairwise independent and that each pair has a BVN distribution.
12. For $(X, Y) \sim$ BVN $(0, 0, 1, 1 ; \rho)$ define a variate V whose values are 1, 2, 3, 4 respectively, depending on which quadrant the observed value for (X, Y) lies in. What is the p.m.f. of V?
13. The height h and weight W of a group of people are approximately normally distributed with mean weight 140 lb, S.D. 12 lb, mean height 68", S.D. 4" and correlation coefficient 2/3. Given that an individual is 66" tall, what is his expected weight?
14. Let (X, Y) be BVN $(0, 0, 1, 1 ; \rho)$. Show that the area of the ellipse

$$(x^2 / \sigma_1^2) - 2\rho(xy / \sigma_1 \sigma_2) + (y^2 / \sigma_2^2) = \lambda^2$$

    of constant probability density is $\pi\lambda^2 \sigma_2 \sigma_2 / \sqrt{1 - \rho^2}$; and hence that the area of strip between the ellipses corresponding to the parameter values λ and $\lambda + d\lambda$ is $2\pi\lambda\, d\lambda\, \sigma_1 \sigma_2 / \sqrt{1 - \rho^2}$. Deduce that the probability that a pair of values (x, y) chosen at random, will be represented by a point inside the strip, is

$$\exp [-\lambda^2 / 2(1 - \rho^2)]\, \lambda\, d\lambda / (1 - \rho^2).$$

Hence by integrating the probability that the point will fall inside the ellipse λ is $1 - \exp[-\lambda^2/2(1-\rho^2)]$.

Also show that, for a given value of $d\lambda$, the probability that the point (x, y) will fall in the strip between the ellipses λ and $\lambda + d\lambda$ is maximum when $\lambda^2 = 1 - \rho^2$.

15. If $(X, Y_0$ is BVN $(10, 20, 16, 100\ ; \rho)$ and $E(Y \mid x = 15) = b$, find k such that $P\{b - k < Y < b + k\} = 0.95$. When (i) $\rho = -0.5$, (ii) $\rho = 0.1$, (iii) $\rho = 0.09$.

16. An object tends to fall on a point target, without systematic error. The distribution about the target is BVN $(0, 0, \sigma^2, \sigma^2\ ; 0)$. Construct a circle and a square, of equal area, centred on the target. Show that the probability is higher for the circle.

17. Let (X, Y) be BVN $(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2\ ; \rho)$. Show that $Z = aX + bY + c$ is $N(a\mu_1 + b\mu_2 + c,\ a^2\sigma_1^2 + 2ab\rho\sigma_1\sigma_2 + b^2\sigma_2^2)$, where a, b, c are constants.

18. Let X_1 and X_2 be i.i.d. $N(0, 1)$ variates and suppose that

$$Y_1 = a_1 + b_{11}X_1 + b_{12}X_2,\quad Y_2 = a_2 + b_{21}X_1 + b_{22}X_2.$$

Show that Y_1 and Y_2 are dependent normal variates and find their variance and covariance.

19. Let X and Y be i.i.d. $N(0, 1)$ variates and suppose that $U = X + Y + Z$, $V = X - Y + 2Z$. Show that U and V have BVN distribution, find Corr (U, V) and $E(U \mid V = 1)$.

20. If (XY) is BVN $(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2\ ; \rho)$, find the necessary and sufficient condition for $X + Y$ and $X - Y$ to be independent.

21. Let (X, Y) be BVN $(0, 0, \sigma_1^2, \sigma_2^2\ ; \rho)$. Find the constants a, b, c, d, s.t. $aX + bY$ & $cX + dY$ and i.i.d. $N(0, 1)$ variates.

22. If $(X, Y) \sim$ BVN $(0, 0, \sigma_1^2, \sigma_2^2\ ; \rho)$, show that $U = X/\sigma_1$, $V = (1 - \rho^2)^{-1/2}\,[(Y/\sigma_2) - (\rho X/\sigma_1)]$ are i.i.d. $N(0, 1)$ variates.

23. Let $(X, Y) \sim$ BVN $(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2\ ; \rho)$. Show that $U = X - \mu_1$, $V = Y - \mu_2 - \rho\sigma_1^{-1}\sigma_2(X - \mu_1)$ are independent normal variates.

24. Unit steps are taken to the right and left (independently) with equal probability. Let S_n denote the position after n steps. Show that the joint p.d.f. of (X, Y) where $X = S_{2n}/\sqrt{n}$ and $Y = (S_{2n} - S_n)/\sqrt{n}$ is asymptotically BVN.

25. Let $(X, Y) \sim$ BVN $(0, 0, \sigma_1^2, \sigma_2^2\ ; \rho)$. Show that rth cumulant, (i.e. semi-invariant) of XY is given by $k_r = \frac{1}{2}(r-1)!\,\sigma_1^r\,\sigma_2^r\,[(\rho + 1)^r + (\rho - 1)^r]$. Deduce that

$$E(X^2Y^2) = \sigma_1^2\,\sigma_2^2\,(1 + 2\rho^2).$$

26. If (X, Y) is BVN $(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2\ ; \rho)$, prove that

$$\mu_{r,s} = \sigma_1^r\,\sigma_2^s \sum_j \binom{s}{j} (1 - \rho^2)^{s-j}\, v_j\, v_{r+s-j} \text{ where } v_k = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^k\, e^{-\frac{1}{2}x^2}\, dx.$$

Hence show that $\mu_{r,s} = 0$, if $r + s$ is odd integer.

27. (a) Let (X_j, Y_j), $1 \le j \le n$ be i.i.d. variates with zero means, variances σ_1^2 and σ_2^2 and correlation coefficient ρ. Show that the limiting distribution of $[(\sigma_1\sqrt{n})^{-1}\,\Sigma X_j, (\sigma_2\sqrt{n})^{-1}\,\Sigma Y_j]$ is standard BVN distribution.

(b) Let $(X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n)$ be a random sample from BVN $(0, 0, 1, 1, \rho)$. Find $E(Z_j)$ where $Z_j = 1$, if $(X_j - \overline{X})(Y_j - \overline{Y}) > 0$; $Z_j = 0$, if $(X_j - \overline{X})(Y_j - \overline{Y}) < 0$.

28. Let (X, Y) by BVN $(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2 ; \rho)$ with $\mu_2 \gg \sigma_2$. Show that to a chose degree of approximation, $P\{(X/Y) < k\} = \Phi(g)$, where $g = (k\mu_2 - \mu_1)/(\sigma_1^2 - 2\rho\sigma_1\sigma_2 + \sigma_2^2)^{1/2}$.

29. Let X and Y be $N(0, 1)$ variates. Show that

$$h(x, y) = f(x)\,g(y)\,\{1 + \alpha\,[2F(x) - 1]\,[2G(y) - 1]\},\ |\alpha| \le 1$$

is a bivariate density function with X and Y as its marginal densities. Show that $Z = X + Y$ is not Normal except in the trivial case $\theta = 0$. Here f, g are the density functions and F, G are the c.d.f's. of X and Y respectively.

30. Given : $\int_0^c \frac{e^{-\frac{1}{2}z^2}}{\sqrt{2\pi}}\,dz = \frac{1}{4}$, ($c = 0.6745$ from Tables). Let

$A_1 = \{(x, y) : x > c, y > 0\}$, $A_2 = \{(x, y) : x < -c, y > 0\}$, $A_3 = \{(x, y), -c < x < c, y < 0\}$

Let $f(x, y) = \pi^{-1} e^{-\frac{1}{2}(x^2 + y^2)}$, $(x, y) \in A_1 \uplus A_2 \uplus A_3$; $f(x, y) = 0$, otherwise.

Show that f is a 2-dim. p.d.f. which is non-Gaussian with Gaussian marginals.

31. Let $f_1(x, y)$ and $f_2(x, y)$ denote respective BVN $(0, 0, 1, 1 ; 0)$ and BNV $(0, 0, 1, 1 ; \rho)$ densities. Suppose (X, Y) have joint density $f(x, y) = \frac{1}{2} f_1(x, y) + \frac{1}{2} f_2(x, y)$.

Show that X and Y have Gaussian marginals, but the joint density is normal iff $\rho = 0$.

The only human Institution which rejects progress is the cemetery.

Science is the belief in the ignorance of the experts. (R.P. Feynman)

Chi-Square Distribution 21

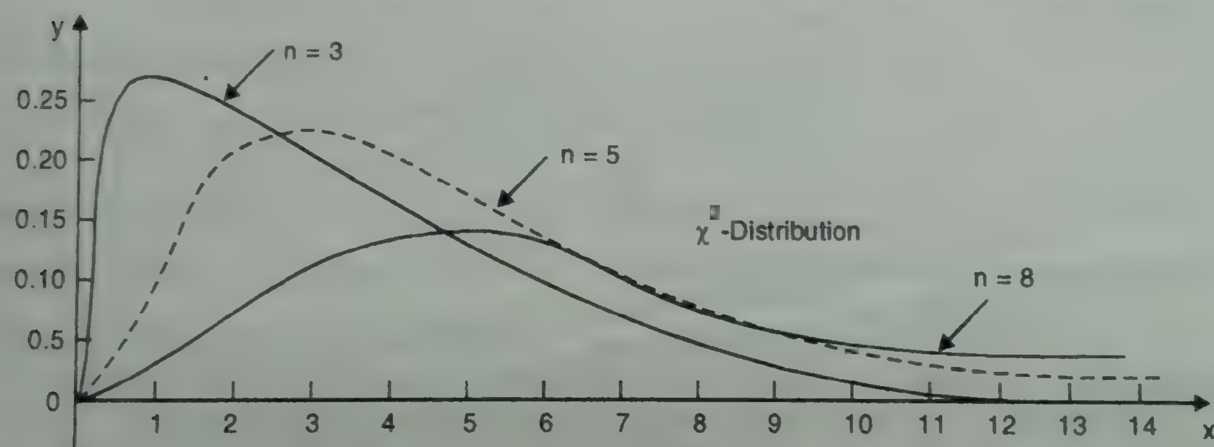
Introduction. In this chapter, we introduce a special case of gamma distribution called **chi-square distribution**, that has earned the distinction of a basic sampling distribution. We present here the theoretical fundamentals, the wide variety of its applications may be considered later on.

Note. Some authors write **Chi-squared distribution** instead of **Chi-square distribution.**

21·10. *Basic Definition*

A random variable X is said to be chi-squared distributed with n degrees of freedom, written $X \sim \chi^2_{(n)}$, if its p.d.f. is given by

$$f(x) = e^{-(x/2)} x^{(n/2)-1}/(2)^{n/2}\,\Gamma(\tfrac{1}{2}n),\quad 0 < x < \infty; \qquad f(x) = 0, \text{ otherwise.}$$



21·11. *Theorem : Sum of Independent Squared Normal Deviates*

If $Z_1, Z_2, ..., Z_n$ are i.i.d. $N(0, 1)$ variates, then $(Z_1^2 + Z_2^2 + ... + Z_n^2) \sim \chi^2_{(n)}$.

[n is the number of independent variates, hence called degrees of freedom (d.f.)]

Proof. If Z is $N(0, 1)$, then the m.g.f. of Z^2 is

$$M(t) = \mathrm{E}(e^{tZ^2}) = \int_{-\infty}^{\infty} e^{tZ^2} \frac{e^{-\frac{1}{2}z^2}}{\sqrt{2\pi}}\,dz = \int_{-\infty}^{\infty} \frac{e^{-(1-2t)z^2}}{\sqrt{2\pi}}\,dz = (1-2t)^{-1/2}\left(\int_{-\infty}^{\infty} \frac{e^{-u^2/2}}{2\pi}\,du\right), [(1-2t)z^2 = u]$$

$$= (1-2t)^{-1/2} \qquad [\because \text{ Area under } N(0, 1) \text{ is unity}] \qquad ...(1)$$

Since $Z_1, Z_2, ..., Z_n$ are i.i.d. it follows that

$$M(t : Z_1^2 + Z_2^2 + ... + Z_n^2) = [M/(t : Z_1^2)]^n = (1-2t)^{-n/2}, \qquad [\text{by } (1)] \qquad ...(2)$$

This shows that $X = \sum Z_i^2$ is gam $(\frac{1}{2}, \frac{1}{2}n)$; hence p.d.f. of X is

$$f(x) = \frac{e^{-x/2} x^{(n/2)-1}}{(2)^{n/2} . \Gamma(\frac{1}{2}n)}, \quad 0 < x < \infty. \qquad [\text{i.e. } X \sim \chi^2_{(n)}] \qquad ...(3)$$

Cor. If X is $\chi^2_{(n)}$, then $Y = \frac{1}{2}X$ is gam $(\frac{1}{2}n, 1) \equiv$ gam $(n/2)$.

Proof. Put $\frac{1}{2}x = y$, $0 < x < \infty \Rightarrow 0 < y < \infty$, $dx = 2dy$. Then (3) gives

$$f_1(y) = \frac{e^{-y}(2y)^{(n/2)-1}\,2}{(2)^{n/2} . \Gamma(n/2)} = \frac{e^{-y}\, y^{(n/2)-1}}{\Gamma(n/2)}, \quad 0 < y < \infty.$$

This show that $Y = \frac{1}{2}X \sim \text{gam}(1, \frac{1}{2}n)$.

Remark. Some authors define $\chi^2_{(n)}$-variate by the relation

$$\chi^2_{(n)} = Z_1^2 + Z_2^2 + ... + Z_n^2. \qquad [Z \sim N_i(0, 1), \text{ i.i.d.}].$$

21-20. *Moments of χ^2-distribution and Recurrence Formula*

$$\mu'_r = E(X^r) = \int_0^\infty x^r \frac{e^{-(x/2)} x^{(n/2)-1}}{\Gamma(\frac{1}{2}n) . (2)^{n/2}}\, dx = \int_0^\infty \frac{e^{-(x/2)}\, x^{r+(n/2)-1}}{(2)^{n/2} . \Gamma(\frac{1}{2}n)}\, dx, \qquad \left[\int_0^\infty e^{-\lambda x} x^{a-1}\, dx = \frac{\Gamma(a)}{\lambda^a}\right]$$

$$= (2)^r\, \Gamma(\tfrac{1}{2}n + r) / \Gamma(\tfrac{1}{2}n). \qquad ...(1)$$

Using $\Gamma(k+1) = k\Gamma(k)$, this gives

$$\mu'_r = \frac{(2)^r}{\Gamma(n/2)} \cdot \left(\frac{1}{2}n + r - 1\right) \Gamma\left(\frac{1}{2}n + r - 1\right)$$

$$= (n + 2r - 2) . \frac{(2)^{r-1} . \Gamma(\frac{1}{2}n + r - 1)}{\Gamma(n/2)}$$

$$\therefore \qquad \mu'_r = (n + 2r - 2)\, \mu'_{r-1}. \qquad ...(1)$$

Obviously, this provides : $\mu'_1 = n$, $\mu'_2 = n(n+2)$; Var $(X) = \mu'_2 - \mu_1'^2 = 2n$.

21-30. *M.G.F. and Pearson's Coefficients*

$$M(t : \chi^2_{(n)}) = (1 - 2t)^{-n/2}, \; M(it : \chi^2_{(n)}) = (1 - 2it)^{-n/2}. \qquad [\S 8\text{-}16(12)]$$

$$K(t : X) = \ln M(t : X) = -\tfrac{1}{2} n \ln(1 - 2t) = \tfrac{1}{2} n \sum_{r=1}^{\infty} \frac{(2t)^r}{r} = \frac{n}{2} \sum_{r=1}^{\infty} 2^r . (r-1)! \frac{t^r}{r!}$$

Thus, $k_r = n 2^{r-1} (r-1)!$. [Coeff. of t^r/r ! in $K(t)$]

Observe $k_1 k_3 = 2k_2^2$

In particular, $k_1 = n$, $k_2 = 2n$, $k_3 = 8n$, $k_4 = 48n$.

Obviously $2\beta_2 - 3\beta_1 - 6 = 0$.

$$M(t) = (1 - 2t)^{-n/2} = \sum_{r=0}^{\infty} \binom{\frac{1}{2}n + r - 1}{r} (2t)^r$$

$$= \sum_{r=0}^{\infty} \frac{\Gamma(\frac{1}{2}n+r)}{\Gamma(n/2)} . (2)^r . \left(\frac{t^r}{r!}\right)$$

This gives : $\mu'_r = 2^r . \Gamma(\frac{1}{2}n+r)/\Gamma(\frac{1}{2}n)$. [§21-20(1)]

Pearson's Coefficients. As $\mu'_1 = k_1 = n, \mu_2 = k_2 = 2n, \mu_3 = k_3 = 8n, \mu_4 = k_4 + 3k_2^2 = 48n + 12n^2$

$$\therefore \beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{64n^2}{8n^3} = \frac{8}{n}; \quad \beta_1 = \frac{\mu_4}{\mu_2^2} = 3 + \frac{12}{n}; \quad \gamma_1 = \sqrt{\beta_1} = \sqrt{\frac{8}{n}}; \quad \gamma_2 = \beta_2 - 3 = \frac{12}{n}.$$

Standardized χ^2-variate : $(\chi^2)^* = (\chi^2 - n)/\sqrt{2n}$.

21-31. *Central-moments Recurrence Formula*

$$\mu_{r+1} = 2r(\mu_r + n\mu_{r-1}), \quad r \geq 1.$$

Proof. $M(t:\chi_n^2) = (1-2t)^{-n/2}$; $\quad M(t:X-n) = e^{-nt}(1-2t)^{-n/2}$.

We denote the central m.g.f by $m(t)$; take logarithmic differentiation w.r.t. "t" to get

$$\ln m(t) = -nt - \tfrac{1}{2}n\ln(1-2t), \quad \frac{m'(t)}{m(t)} = -n + \frac{n}{1-2t} = \frac{2nt}{1-2t}.$$

$$\therefore \quad (1-2t)\, m'(t) = 2nt\, m(t)$$

We differentiate this equation r times w.r.t. "t". use Leibnitz Rule to get

$$(1-2t)[m(t)]^{(r+1)} + r(-2)\, m(t)^{(r)} = 2nt\,[m(t)]^{(r)} + r\,.\,2n[m(t)]^{(r-1)}$$

Putting $t = 0$, using $m^{(r)}(0) = \mu_r$; we recover : $\mu_{r+1} = 2r(\mu_r + n\mu_{r-1})$.

Example : Let $Y = X^2$ where X is $N(0, 1)$. Calculate $\mathsf{E}(Y)$ by two different ways.

Solution. $\mathsf{E}(Y) = \mathsf{E}(X^2) = \mathrm{Var}(X) = 1, \quad [X \sim N(0,1)]$

$\mathsf{E}(Y) = \mathsf{E}[\chi^2_{(1)}]$ = Mean of chi-square with 1 degree of freedom = 1.

Remark. We need not use $f_X(x)$ and $f_Y(y)$, since the results are too well-known.

21-32. *Reproductive (or Additive) Property*

If $X \sim \chi^2_{(m)}$ and $Y \sim \chi^2_{(n)}$ are independent variates, then $X + Y \sim \chi^2_{(m+n)}$.

Proof. Since X and Y are independent variates

$$M(t:X+Y) = M(t:X)\,M(t:Y) = (1-2t)^{-m/2} . (1-2t)^{-n/2} = (1-2t)^{-(m+n)/2}$$

This shows that $(X + Y)$ is $\chi^2_{(m+n)}$.

Extension to n independent χ^2-variates is obvious.

Cor 1. If $X \sim \chi^2_{(m)}, Y \sim \chi^2_{(n)}$ are independent variates, that $(Y - X) \sim \chi^2_{n-m}$, provided X and $Y - X$ are independent $(n > m)$.

Proof. Write $Y = X + (Y - X)$; then since X and $Y - X$ independent.

$$M(t:Y) = M[t:X+(Y-X)] = M(t:X)\,M(t:Y-X)$$

$$\therefore \quad (1-2t)^{-n/2} = (1-2t)^{-m/2}.M(t:Y-X) \Rightarrow M(t:Y-X) = (1-2t)^{-(n-m)/2}.$$

This shows that $(Y-X) \sim \chi^2_{(n-m)}$.

Cor. 2. If X is $\chi^2_{(m)}$, $X+Y$ is $\chi^2_{(m+n)}$ where X and Y are independent, then Y is $\chi^2_{(n)}$.

Proof.
$$M(t:X+Y) = M(t:X)\,M(t:Y) \qquad (\because X \;\&\; Y \text{ are indep.})$$

$$\therefore \quad (1-2t)^{-(m+n)/2} = (1-2t)^{-m/2}.M(t:Y) \Rightarrow M(t:Y) = (1-2t)^{-n/2}.$$

This shows that Y is $\chi^2_{(n)}$.

Remark. Converse to reproductive property holds in the following sense :

If $X \sim \chi^2_{(m)}, Y \sim \chi^2_{(n)}, (X+Y) \sim \chi^2_{(m+n)}$, then X & Y are independent.

21-33. *Fisher's Approximation*

For large values of n, $\qquad \chi^2_{(n)} \rightarrow N(n, 2n)$.

Proof. Let $Y = (\chi^2 - n)/\sqrt{2n}$ be standardized χ^2-variate. Since $\phi(t:\chi^2) = (1-2it)^{-n/2}$, hence

$$\phi(t:Y) = \phi[t:(\chi^2-n)/\sqrt{2n}] = e^{-it\sqrt{n/2}}\,\phi(t/\sqrt{2n}:\chi^2) = e^{-it\sqrt{n/2}}\,[1-(2it/\sqrt{2n}]^{-n/2}.$$

Taking logarithms, we obtain

$$\ell n\,\phi(t:Y) = -it\sqrt{n/2} - \tfrac{1}{2}n\,\ell n[1 - it(2/n)^{1/2}] \qquad [\text{Assume } \sqrt{2}\,t < n]$$

$$= -it\sqrt{n/2} + \tfrac{1}{2}n\sum_{r=1}^{\infty}\frac{(it)^r}{r}\left(\frac{2}{n}\right)^{r/2} = -\frac{1}{2}t^2 + \frac{1}{2}n\sum_{r=3}^{\infty}\frac{(it)^r}{r}\left(\frac{2}{n}\right)^{r/2}.$$

As $n \rightarrow \infty$, all the terms in the infinite series vanish. Hence using continuity of ℓn function

$$\lim_{n\to\infty} \ell n\,\phi(t:Y) = -\tfrac{1}{2}t^2 \Rightarrow \lim_{n\to\infty}\phi(t:Y) = e^{-\frac{1}{2}t^2}.$$

This shows that when n is large, $Y = \chi^{2*}$ is $N(0, 1)$ or equivalently χ^2 is $N(n, 2n)$.

Remark. That expectation and variance in the approximating normal variate depends on n is of great disadvantage.

Aliter. Let $Y_j \sim \chi^2_{(1)}$, $1 \le j \le n$ be independent variates. By Addition theorem $S_n = Y_1 + \ldots + Y_n$ is χ^2 with n degrees of freedom. Since $E(\chi^2) = n, \text{Var}(\chi^2) = 2n$, by C.L.T. $S^*_n = (\chi^2 - n)/\sqrt{2n}$ converges weakly to $N(0, 1)$, which is equivalent to the fact that χ^2 is $N(n, 2n)$ for large n.

21-40. *Mode. Points of Inflexion. Pearson's Coefficient of Skewness*

We set $K = [2^{n/2}\,\Gamma(\frac{1}{2}n)]^{-1}$; then p.d.f. of $\chi^2_{(n)}$ is $y = Ke^{-x/2}\,.\,x^{(n/2)-1}$, $0 < x < \infty$

$$\therefore \quad \ln y = \ln K - \tfrac{1}{2}x + (\tfrac{1}{2}n - 1)\ln x.$$

Differentiating w.r.t. "x", writing $y' = (dy/dx)$, etc. we get

$$(y'/y) = -\tfrac{1}{2} + (\tfrac{1}{2}n - 1)\,x^{-1}\,; (y''/y) - (y'/y)^2 = -(\tfrac{1}{2}n - 1)\,x^{-2}. \qquad \ldots(1)$$

Putting $y' = 0$, we get $x = n - 2$. And at $x = n - 2$, $n > 2$, $y'' < 0$.

Hence the modal value of a $\chi^2_{(n)}$-variate is at $x = n - 2$, $n > 2$.

Inflectional Points. Now put $y'' = 0$ and eliminate (y'/y) between 1(a) and 1(b) to get

$$\frac{n-2}{2x^2} = \left(\frac{n-2}{2x} - \frac{1}{2}\right)^2 \Rightarrow [x - (n-2)]^2 = 2(n-2) \Rightarrow x = (n-2) \pm [2(n-2)]^{1/2}.$$

Thus, the points of inflexion are equidistant from the mode.

Skewness.
$$S_k = \frac{\text{Mean} - \text{Mode}}{\text{S.D.}} = \frac{n - (n-2)}{\sqrt{2n}} = \left(\frac{2}{n}\right)^{1/2}$$

Since $S_k > 0$, it follows that $\chi^2_{(n)}$ is positively skewed.

21-41. *Mean Absolute Deviation*

For $\chi^2_{(n)}$, $f(x) = Ke^{-x/2}x^{(m/2)-1}$, $[K^{-1} = 2^{n/2}\,.\,\Gamma(\frac{1}{2}n)]$, $x > 0$ and $E(X) = n$. ...(1)

$$\therefore \quad M \equiv E(|X - n|) = \int_{x<n}(n-x)\,f(x)\,dx + \int_{x \ge n}(x-n)\,f(x)\,dx \qquad \ldots(i)$$

$$0 \equiv E(X - n) = \int_{x<n}(x-n)\,f(x)\,dx + \int_{x \ge n}(x-n)\,f(x)\,dx \qquad \ldots(ii)$$

Subtracting (ii) from (i) gives

$$M = 2\int_{x<n}(n-x)\,f(x)\,dx = 2n\int_0^n f(x)\,dx - 2\int_0^n x\,f(x)\,dx \qquad \ldots(2)$$

We now consider the second integral, integrate it by parts and get

$$\int_0^n xf(x)\,dx = K\int_0^n x^{n/2}\,e^{-x/2}dx = -2Ke^{-n/2}(n)^{n/2} + nK\int_0^n x^{(n/2)-1}\,e^{-x/2}\,dx$$

$$= -2Ke^{-n/2}.(n)^{n/2} + n\int_0^n f(x)\,dx, \qquad \text{[by (1)]}$$

Substituting this expression into (2), the two integrals involved therein cancel out, and we are left with

$$M = 4Ke^{-n/2}.(n)^{n/2} = e^{-n/2}(n)^{n/2}/2^{(n-4)/2}\,\Gamma(\tfrac{1}{2}n).$$

21-42. Worked-out Problems

Example* 1. *(Pearson's P_λ-Statistic or P_λ-Test). If X is U(0, 1), show that $Y = -2 \ln X$ is $\chi^2_{(2)}$. Show further that if $X_1, X_2, ..., X_n$ are i.i.d. $U(0, 1)$ variates, and $P = X_1 . X_2 . ... X_n$ (product of variables) then $-2 \ln_e P$ is $\chi^2_{(2n)}$.

Solution. $$M(t:Y) = E(e^{tY}) = E(e^{-2t \ln X}) = E(X^{-2t}) = \int_0^1 x^{-2t} dx = (1-2t)^{-1}. \quad ...(i)$$

This shows that Y is $\chi^2_{(2)}$. Further

$$-2\ln P = -2\ln(X_1 X_2 ... X_n) = Y_1 + Y_2 + ... + Y_n; \quad [Y_j = -2\ln X_j].$$

$$M(t: -2\ln P) = M(t: Y_1 + Y_2 + ... + Y_n) = [M(t: Y_j)]^n = (1-2t)^{-n}. \quad \text{[by (i)]}$$

This shows that $-2 \ln P \sim \chi^2_{(2n)}$.

Note. The P_λ-Statistic is useful in combining several independent significance Tests of Statistical hypothesis.

***Example* 2.** If X and Y are i.i.d. Unif(0, 1) variates, show that

$$U = (-2\ln X)^{(1/2)} \cos 2\pi Y, \quad V = (-2\ln X)^{1/2} \sin 2\pi Y \qquad \textbf{[Box-Muller Transformations]}$$

and i.i.d. $N(0, 1)$ variates. Deduce that U^2 and V^2 are i.i.d. $\chi^2_{(1)}$, variates.

Solution. Let $u = (-2\ln x)^{1/2} \cos 2\pi y, \qquad v = (-2\ln x)^{1/2} \sin 2\pi y. \quad ...(1)$

$$J^{-1} = \frac{\partial(u,v)}{\partial(x,y)} = \begin{vmatrix} (-2/\ln x)^{1/2} \dfrac{\cos 2\pi y}{2x} & -2\pi(-\ln x)^{1/2} \sin 2\pi y \\ (-2/\ln x)^{1/2} \dfrac{\sin 2\pi y}{2x} & 2\pi(-2\ln x)^{1/2} \cos 2\pi y \end{vmatrix} = -\frac{2\pi}{x}$$

Thus, $dx\, dy = | x/2\pi |\, du\, dv$. From (1), we also get

$$u^2 + v^2 = -2\ln x \Rightarrow x = \exp[-\tfrac{1}{2}(u^2 + v^2)].$$

Since X and Y are i.i.d. $U(0, 1)$, their joint elemental p.d.f. is $dP_1(x, y) = 1\, dx\, dy$ which under the transformation (1) becomes

$dP_2(u, v) = (x/2\pi)\, du\, dv = (2\pi)^{-1} \exp[-\tfrac{1}{2}(u^2 + v^2)]\, du\, dv$, $-\infty < u, \ v < \infty$. This is the joint p.d.f. of two indep. $N(0, 1)$ variates. That is $U \sim N(0, 1)$ and $V \sim N(0, 1)$ are indep. variates. Since U and V are indep. so U^2 and V^2 must be independent. Further, $[N(0, 1)]^2 \sim \chi^2_{(1)}$ so $U^2 \sim \chi^2_{(1)}$ and $V^2 \sim \chi^2_{(1)}$ are i.i.d.

***Example* 3.** If $X \sim \chi^2_{2m}$ and $Y \sim \chi^2_{2n}$ are indep. variates, find the p.d.f. of $U = aX + bY$, $(a > 0, b > 0)$.

Solution. The joint c.d.f. differential of X and Y is

$$dF(x, y) = \frac{e^{-x/2} . x^{m-1}\, dx}{\Gamma(m)\, 2^m} . \frac{e^{-y/2}\, y^{n-1}\, dy}{\Gamma(n) . 2^n} \quad ...(1)$$

Let $u = ax + by$, $v = ax/u$, so that $ax = uv$, $by = u(1-v)$, $0 \le u < \infty$, $0 < v < 1$.

$$\frac{\partial(x, y)}{\partial(u, v)} = \frac{1}{ab}\begin{vmatrix} v & u \\ 1-v & -u \end{vmatrix} = \frac{-u}{ab},\ dx\,dy = \frac{|u|}{ab}\,du\,dv.$$

The joint c.d.f. differential of U and V, using (1), is

$$dF(u, v) = (Ke^{-u/2b}\, u^{m+n-1}\, du).\ (v^{m-1}(1-v)^{n-1}\, e^{u(1-c)v/2b}\, dv) \qquad ...(2)$$

where $K^{-1} = a^m b^n 2^{m+n}\,\Gamma(m)\,\Gamma(n)$, and $c = (b/a)$.

To obtain the p.d.f of U, we integrate out the unwanted variable v in (2). For this, we use exponential series $e^{\theta} = \Sigma(\theta^r/r!)$, assume term by term integration to infinite series and obtain

$$\int_0^1 v^{m-1}(1-v)^{n-1}\sum_{r=0}^{\infty}\frac{(u/2b)^r\,(1-c)^r\,.\,v^r}{r!} = \sum_{r=0}^{\infty}\left(\frac{u}{2b}\right)^r\frac{(1-c)^r}{r!}\int_0^1 v^{m+r-1}(1-v)^{n-1}\,dv$$

$$= \sum_{r=0}^{\infty}\left(\frac{u}{2b}\right)^r(1-c)^r\,\frac{1}{r!}\,\frac{\Gamma(m+r)\,\Gamma(n)}{\Gamma(m+n+r)} \qquad ...(3)$$

Inserting (3) into integration w.r.t. v of Eq. (2) we get $[\lambda = m + n + r]$

$$\frac{dF(u)}{du} = \left(\frac{b}{a}\right)^m\frac{1}{\Gamma(m)}\sum_{r=0}^{\infty}\left(\frac{u}{b}\right)^{r+m+n-1}\frac{e^{-u/2b}(1-c)^r\,\Gamma(m+r)}{2^{\lambda}\,\Gamma(\lambda)\,r!},\ \left[c = \frac{b}{a}\right].$$

$$f_U(u) = \frac{c^m}{2\Gamma(m)}\cdot\sum_{r=0}^{\infty}\frac{\Gamma(m+r)}{r!}(1-c)^r\,\frac{e^{-u/2b}.\,(u/2b)^{\lambda-1}}{\Gamma(\lambda)}. \qquad [\lambda = m + n + r]$$

Note. $M(t : u) = M(t : aX + bY) = M(at : X)\,M(bt : Y) = (1-2at)^{-m}\,(1-2bt)^{-n}$

$$= \sum_{r=0}^{\infty}\sum_{s=0}^{\infty}\binom{m+r-1}{r}\binom{n+s-1}{s}(2at)^r\,(2bt)^s. \qquad \text{Simplify and identify.}$$

Problems with Solutions Provided at the End of the Text

1*. Show that for 2 d.f. the probability P_0 of a value of $\chi^2 > \chi_0^2$ is $\exp(-\frac{1}{2}\chi_0^2)$ and hence that $\chi_0^2 = 2\ln(1/P_0)$. Deduce the value of χ_0^2 when $P_0 = 0.05$.

2*. If $X = 2$ is the unique mode of gam $(2, \lambda)$ distribution, find the parameter λ and evalutate $P\{X < 9.49\}$.

3*. If X is $\chi^2_{(2N)}$ and N is geom (p), find the unconditional m.g.f. of X.

21-50. Chi Distribution

Definition. The positive square-root of χ^2 with n, d.f. is called the *Chi-distribution* with n, d.f. and is written $\chi_{(n)}$. Thus $\chi_{(n)} = \sqrt{\chi^2_{(n)}}$.

Density Function. $F(x) = P\{\chi_{(n)} \le x\} = P\{\chi^2_{(n)} \le x^2\} = \int_0^{x^2} \dfrac{e^{-y/2}\, y^{(n/2)-1}\, dy}{\Gamma(n/2).(2)^{n/2}}$.

Differentiating under integral sign (DUIS), using $f(x) = F'(x)$, we get

$$f(x) = e^{-x^2/2}(x^2)^{(n/2)-1}.2x/(2)^{n/2}\,\Gamma(n/2), \quad x>0$$

$$f(x) = 2x^{n-1}e^{-x^2/2}/\Gamma(\tfrac{1}{2}n)(2)^{n/2}, \quad x>0; \qquad f(x) = 0, \text{ otherwise.}$$

Simple Moments :

$$E(X^r) = \frac{2}{\Gamma(\frac{1}{2}n).(2)^{n/2}}\int_0^\infty x^{r+n-1}\,e^{-x^2/2}\,dx = \frac{(2)^{r/2}}{\Gamma(\frac{1}{2}n)}\int_0^\infty z^{(1/2)(n+r)-1}e^{-z}\,dz. \quad [z = \tfrac{1}{2}x^2]$$

$\therefore \quad \mu_r' = (2)^{r/2}\,\Gamma[\frac{1}{2}(n+r)]/[\Gamma(\frac{1}{2}n) = (2)^{r/2}\,(\frac{1}{2}n)^{[r/2]}$ [Reverse factorial power]

Thus, $\quad E(\chi_n) = \sqrt{n}, \qquad \mu_2' = E(\chi_n^2) = n$ (mean value of χ_n^2).

Example : Show that, for large values of n, $E(\chi_n^2) = [E(\chi_n)]^2$.

Solution. We know that, $\mu_r' = (\sqrt{2})^r\,(\frac{1}{2}n)^{[r/2]}$. So, $\mu_1' = E(\chi_{(n)}) = \sqrt{n} \Rightarrow [E(\chi_{(n)})]^2 = n = E(\chi^2_{(n)})$.

21-60. *Quotient Theorem : Ratio of Two Independent Chi-square Variates*

If $X \sim \chi^2_{(m)}$ and $Y \sim \chi^2_{(n)}$ are independent variates, then $(X/Y) \sim B_{II}\,(\frac{1}{2}m, \frac{1}{2}n)$.

Proof. Recall : $\chi^2_{(n)} = \text{gam}\,(\frac{1}{2}n, \frac{1}{2}), \chi^2_{(m)} = \text{gam}\,(\frac{1}{2}n, \frac{1}{2})$.

By § 19-33, [gam (a, λ)/gam (b, λ)] = $B_{II}\,(a, b)$.

Thus, $\quad \chi^2_{(m)}/\chi^2_{(n)} = B_{II}\,(\frac{1}{2}m, \frac{1}{2}n)$

(Reproduce the derivation)

Example 1. If $X \sim N(\mu_1, \sigma_1^2)$ and $Y \sim N(\mu_2, \sigma_2^2)$ are independent, derive the distribution of $Z = (X-\mu_1)/(Y-\mu_2)$.

Solution.
$$Z^2 = \left[\frac{(X-\mu_1)/\sigma_1}{(Y-\mu_2)/\sigma_2}\right]^2 \cdot \left(\frac{\sigma_1^2}{\sigma_2^2}\right), \quad \Rightarrow \quad \frac{\sigma_2^2}{\sigma_1^2}Z^2 = \frac{Z_1^2}{Z_2^2} = \frac{[N(0,1)]^2}{[N(0,1)]^2}$$

where Z_i is $N(0, 1)$ or equivalently Z_i^2 is $\chi^2_{(1)}$. Thus, $(\sigma_2^2/\sigma_1^2)\,Z^2$ is the quotient of two independent χ^2 variates and hence is a $B_{II}\,(\frac{1}{2}, \frac{1}{2})$ variate. Consequently, the probability differential is

$$dF(u) = \frac{1}{B(\frac{1}{2},\frac{1}{2})}\cdot\frac{(\sigma_2^2z^2/\sigma_1^2)^{1/2-1}\,du}{[1+(\sigma_2^2z^2/\sigma_1^2)]^{1/2+1/2}} = \frac{\sigma_1\sigma_2}{\pi(\sigma_1^2+z^2\sigma_2^2)}\,\frac{dz^2}{z}, \qquad \left[u = \frac{\sigma_2^2}{\sigma_1^2}z^2,\ B(\tfrac{1}{2},\tfrac{1}{2}) = \pi\right]$$

$$\therefore \quad dF(z) = \frac{\sigma_1\sigma_2\,dz}{\pi(\sigma_1^2+\sigma_2^2z^2)}, \quad -\infty < z < \infty, \qquad [\text{Cauchy } (0, \sigma_1/\sigma_2)]$$

Aliter. See Example 16-27.

Example 2. Let X, Y, Z be independent χ^2-variables with $2a$, $2b$, $2c$ degrees of freedom. Prove that $U = \dfrac{X/2a}{Y/2b}$, $V = \dfrac{Z/2c}{(X+Y)/(2a+2b)}$ are independent F-variables.

Solution. Recall : If $X \sim F(2m, 2n)$, then its p.d.f. is

$$f(x) = \frac{m}{n} \cdot \frac{1}{B(m,n)} \cdot \frac{(mx/n)^{m-1}}{[1+(mx/n)]^{m+n}} \equiv \frac{(n/m)^n}{B(m,n)} \cdot \frac{x^{m-1}}{[x+(n/m)]^{m+n}}, 0 < x < \infty \qquad \ldots(1)$$

Now the joint p.d.f. of X, Y, Z is

$$dF(x, y, z) = \frac{e^{-(x+y+z)/2}\, x^{a-1}\, y^{b-1}\, z^{c-1}\, dx\, dy\, dz}{\Gamma(a)\,\Gamma(b)\,\Gamma(c)\,(2)^{a+b+c}}, \quad 0 < x, y, z < \infty \qquad \ldots(2)$$

Put $\quad u = \dfrac{b}{a}\dfrac{x}{y} \equiv \theta \dfrac{y}{x}$, $v = \dfrac{a+b}{c}\dfrac{z}{x+y} \equiv \dfrac{\lambda z}{x+y}$, $w = z.$ $\left[\theta = \dfrac{b}{a},\ \lambda = \dfrac{a+b}{c}\right]$

Then $x = \dfrac{\lambda u w}{v(u+\theta)}$, $y = \dfrac{\lambda\theta w}{v(u+\theta)}$, $z = w$, $x+y+z = w\left(1+\dfrac{\lambda}{v}\right)$, $J^{-1} = \dfrac{\partial(u, v, w)}{\partial(x, y, z)} = \dfrac{v\theta}{y^2}$

The joint distribution differential for new variables is

$$dF(u, v, w) = \frac{e^{-w[1+(\lambda/v)]/2}}{\Gamma(a)\,\Gamma(b)\,\Gamma(c)\,2^{a+b+c}} \left[\frac{\lambda uw}{v(u+\theta)}\right]^{a-1} \left[\frac{\lambda\theta w}{v(u+\theta)}\right]^{b+1} \frac{w^{c-1}}{v\theta}\, du\, dv\, dw \qquad \ldots(3)$$

We integrate **out** the unwanted variable w in (3)

$$\int_0^\infty e^{-w[1+(\lambda/v)]/2}\, w^{a+b+c-1}\, dw = \frac{\Gamma(a+b+c)\,(2)^{a+b+c}}{[1+(\lambda/v)]^{a+b+c}} \qquad \ldots(4)$$

Using (4) into integral of (3) w.r.t. w we recover

$$dF(u, v) = \left(\frac{\theta^b}{B(a,b)} \frac{u^{a-1}\, du}{(\theta+u)}\right)\left(\frac{\lambda^{a+b}}{B(a+b,c)} \frac{v^{c-1}\, dv}{(\lambda+v)^{a+b+c}}\right), \quad 0 < u,\ v < \infty. \qquad \ldots(5)$$

where $B(l, m) = \Gamma(l)\,\Gamma(m)/\Gamma(l+m)$.

Result (5) implies that U and V are indep. variables (by Factor theorem) with F-distributions as guaranteed by (1). [See §22-40]

21-61. *Sum and Quotient of Independent Chi-square Variates*

If $X \sim \chi^2_{(m)}$ and $Y \sim \chi^2_{(n)}$ are independent, then $U = X + Y$, $V = X/(X+Y)$ are independent $\chi^2_{(m+n)}$ and $B_I\left(\frac{1}{2}m, \frac{1}{2}n\right)$ variates.

Recall : $\chi^2_{(m)} = \text{gam}\left(\frac{1}{2}m, \frac{1}{2}\right)$, $\quad \chi^2_{(n)} = \text{gam}\left(\frac{1}{2}n, \frac{1}{2}\right)$. $\qquad$ [Exploit §19-25 Example 1]

21-70. *Chi-square in Terms of Sample Variance*

Let $X_1, X_2, \ldots, X_n$ be a random sample from $N(\mu, \sigma^2)$. Let $nS^2 = \Sigma(X_j - \overline{X})^2$, $1 \le j \le n$. Then

$$nS^2/\sigma^2 \sim \chi^2_{(n-1)}.$$

Proof. Observe : $\bar{X} = (\Sigma X_j)/n \Rightarrow \Sigma(X_j - \bar{X}) = 0, \; 1 \le j \le 1$ (i)

$\therefore \; \Sigma(X_j - \mu)^2 = \Sigma[(X_j - \bar{X}) + (\bar{X} - \mu)]^2 = [\Sigma(X_j - \bar{X})^2] + n(\bar{X} - \mu)^2$, [use (i)] (ii)

Dividing (ii) by σ^2, putting $nS^2 = \Sigma(X_j - \bar{X})^2$, we obtain

$$\sum_{j=1}^{n}\left(\frac{X_j - \mu}{\sigma}\right)^2 = \frac{nS^2}{\sigma^2} + \left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}}\right)^2 \text{ or } W^2 = \left(\frac{nS^2}{\sigma^2}\right) + Z^2 \text{ (with obvious notations)} \quad (1)$$

The term W^2 on the L.H.S. of Eq. (1) is $\chi^2_{(n)}$, being the sum of n indep. $N(0, 1)$ squared variates. Also $Z = (\bar{X} - \mu)/(\sigma/\sqrt{n}) \sim N(0,1)$ so that Z^2 is $\chi^2_{(1)}$. Further, nS^2/σ^2 and $n(\bar{X} - \mu)^2/\sigma^2$ are independently distributed variates.

$$\therefore \quad M(t : W^2) = M(t : (nS^2/\sigma^2) + Z^2) = M(t : nS^2/\sigma^2) \, . \, M(t : Z^2)$$

i.e. $$(1 - 2t)^{-n/2} = M(t : nS^2/\sigma^2) \, . \, (1 - 2t)^{-1/2} \Rightarrow M(t : nS^2/\sigma^2) = (1 - 2t)^{-(n-1)/2}.$$

This proves that $(nS^2/\sigma^2) \sim \chi^2_{(n-1)}$.

Note. $nS^2 = (n - 1)\,\hat{S}^2 = \Sigma(X_i - \bar{X})^2, 1 \le i \le n$. Thus $(v\hat{S}^2/\sigma^2) \sim \chi^2_{(n-1)}$ where $v = (n - 1)$.

21-71. *Distribution of Sample Variance S^2*

$$(1 - 2t)^{-(n-1)/2} = M\left(t : \frac{nS^2}{\sigma^2}\right) = M\left(\frac{nt}{\sigma^2} : S^2\right) = M(\theta : S^2) = \left(1 - \frac{2\sigma^2}{n}\theta\right)^{-(n-1)/2}$$

This shows that $S^2 \sim$ gam $[n/2\sigma^2, (n - 1)/2]$. The p.d.f. of $Y = S^2$ is thus

$$f(y) = \left(\frac{n}{2\sigma^2}\right)^{(n-1)/2} \frac{e^{-ny/2\sigma^2} \, . \, (y)^{(n-3)/2}}{\Gamma[(n-1)/2]}, \quad 0 < y < \infty.$$

Cor. $M(t : v\hat{S}^2/\sigma^2) \, . \, M(vt/\sigma : \hat{S}^2) = (1 - 2t)^{-(n-1)/2} \Rightarrow M(\theta : \hat{S}^2) = (1 - 2\sigma\theta/v)^{-(n-1)/2}.$

This shows that $\hat{S}^2 \sim \text{gam}\,[v/2\sigma^2, (n - 1)/2]$. The p.d.f. of $Z = \hat{S}^2$ in thus

$$f(z) = \left(\frac{n-1}{2\sigma^2}\right)^{(n-1)/2} \cdot \frac{\exp[-(n-1)\,z/2\sigma^2] \, . \, z^{(n-3)/2}}{\Gamma[(n-1)/2]}, \quad 0 < z < \infty.$$

21.72. *Worked-out Problem*

Example : Let $X_i \sim N(\mu, \sigma_i^2)$, $1 \le i \le n$ be indep. $U = K^{-1}\Sigma(X_i/\sigma_i^2)$ and $V = \Sigma(X_i - U)^2/\sigma_i^2$, where $K = \Sigma(1/\sigma_i^2)$ are independently distributed. Show that U is Normal and V is $\chi^2_{(n-1)}$.

Solution. Since U is a linear combination of normal variates, U itself must be normal. Also

$$\text{E}(U) = K^{-1}\Sigma\,[\text{E}(X_i)/\sigma_i^2] = K^{-1}\,\mu\,\Sigma(1/\sigma_i^2) = \mu.$$

$$\text{Var}(U) = K^{-2}\,\Sigma[\text{Var}(X_i)/\sigma_i^4] = K^{-2}\,\Sigma(1/\sigma_i^2) = K^{-1}.$$

So $U \sim N(\mu, K^{-1})$, or $\sqrt{K}\,(U-\mu) \sim N(0,1)$, i.e. $W = K(U-\mu)^2 \sim \chi^2_{(1)}$. ...(1)

$$V = \sum_{i=1}^{n} \frac{[(X_i-\mu)-(U-\mu)]^2}{\sigma_i^2} = \sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma_i}\right)^2 + (U-\mu)^2\,\Sigma\left(\frac{1}{\sigma_i^2}\right) - 2(U-\mu)\sum_{i=1}^{n}\frac{(X_i-\mu)}{\sigma_i^2}$$

$$= \Sigma Z_i^2 + K(U-\mu)^2 - [2(U-\mu)[KU-\mu K] = Z - K(U-\mu)^2]. \quad [Z = \Sigma Z_i^2] \qquad ...(2)$$

$\therefore\ V = Z - W$, where $Z \sim \chi^2_{(n)}, W \sim \chi^2_{(1)}$, by (1) and (2).

So $M(t : V+W) = M(t : Z) \Rightarrow M(t : V).(1-2t)^{-1/2} = (1-2t)^{-n/2}$. [$V$ & W are indep.]

Thus $\qquad M(t : V) = (1-2t)^{-(n-1)/2} \Rightarrow V \sim \chi^2_{(n-1)}$.

Problems with Solutions Provided at the End of the Text

1*. A random sample $X_1, X_2, \ldots, X_n$ of size n is selected from $N(\mu, 1)$ population. Later an additional independent observation X_{n+1} is obtained from the same population. Find the distribution of $(X_{n+1}-\mu)^2 + \sum_{i=1}^{n}(X_i - \overline{X})^2$.

2*. If $X_1, X_2, \ldots, X_n$ is a random sample of X, show that $\mathsf{E}(\hat{S}^2) = \sigma^2$. Does this imply that $\mathsf{E}(\hat{S}) = \sigma$, when X is $N(\mu, \sigma^2)$?

3*. Three samples of sizes n_1, n_2, n_3 are drawn from $N(\mu_i, \sigma_i^2)$, $i = 1, 2, 3$. Let $U = \hat{S}_1^2/\hat{S}_3^2$, $V = \hat{S}_2^2/\hat{S}_3^2$ whre $\hat{S}_i^2$ are sample variances. Find the joint density of U and V.

21-73. *Matrix-grid Theorem*

Let $X_1, X_2, \ldots, X_n$ be independent $N(0, 1)$ variates. Suppose

$$b_{i1}X_1 + b_{i2}X_2 + \ldots + b_{in}X_n = 0, \quad i = 1, 2, \ldots, m. \qquad ...(1)$$

The distribution of $X_1^2 + X_2^2 + \ldots + X_n^2$, subject to $m(< n)$ independent linear constraints (1) is a χ^2 distribution with $n - m$ degrees of freedom.

Proof. The coefficient matrix B of m constraints is of order $m \times n$ and hence its rank is m (it has m independent rows). By Gram-Schmidt Orthogonalization Process, the matrix equation $BX = 0$ can be transformed to $AX = 0$ where $A_{m\times n}$ consists of m unitary, mutually orthogonal vectors $a_i = (a_{i1}, a_{i2}, \ldots, a_{in})$, i = 1, 2, .., m. Thus

$$a_{i1}X_1 + a_{i1}X_2 + \ldots + a_{in}X_n = 0, \quad i = 1, 2, \ldots, m. \qquad ...(2)$$

Let us transform the variates X_i to Y_i, $1 \le i \le n$ by means of linear orthogonal transformations $Y = AX$, where $A = [a_{ij}]$ is $n \times n$ orthogonal matrix. In this system of n equations, $Y_1, Y_2, \ldots, Y_m$ are each zero by (2). Hence

$$\sum_{i=1}^{n} X_i^2 = \sum_{i=1}^{m} Y_i^2 + \sum_{i=m+1}^{n} Y_i^2. \qquad ...(3)$$

By Fisher's Lemma, Y_i, $i = 1, 2, ..., n$ are i.i.d. $N(0, 1)$ variates, hence the distribution of $\sum X_i^2$ subject to constraints (1) is the same as the distribution of $Y_{m+1}^2 + Y_{m+2}^2 + ... + Y_n^2$, by (3), which is obviously $\chi^2_{(n-m)}$.

Example : Let $X_1, X_2, ..., X_n$ be i.i.d. $N(0, \sigma^2)$ variates and

$$Y_i = a_{i1}X_1 + a_{i2}X_2 + ...+ a_{in}X_n, \quad i = 1, 2, ..., n$$

where $\Sigma a_{ij}\, a_{kj} = \delta_{ik}\ (j = 1, 2, ..., n); \delta_{ik}$ being Kroneker delta. Show that

$$[(X_1^2 + ... + X_n^2) - (Y_1^2 + ... + Y_m^2)] / \sigma^2 \sim \chi^2_{(n-m)}.$$

Solution. Since $\delta_{ik} = 0$, if $i \neq k$ and $\delta_{ik} = 1$, if $i = k$, it follows that the matrix $A = [a_{ij}]$ is orthogonal and hence $Y = AX$ is linear orthogonal mapping. Consequently.

$$X_1^2 + X_2^2 + ... + X_n^2 = Y_1^2 + Y_2^2 + ... + Y_n^2$$

So $[(X_1^2 + X_2^2 + ... + X_n^2) - (Y_1^2 + Y_2^2 + ... + Y_m^2) = Y_{m+1}^2 + Y_{m+2}^2 + ... + Y_n^2.$

By Fisher's Lemma $Y_i \sim N(0, \sigma^2)$ are independent variates, hence $(Y_i/\sigma) \sim N(0, 1)$. Consequently

$$(Y_{m+1}^2 + ... + Y_n^2) / \sigma^2 \sim \chi^2_{(n-m)}. \qquad \text{[Sum of } (n - m) \text{ i.i.d. } \chi^2 \text{ variates]}.$$

21-74. *Distn. of Sample Corr. Coeff. when Population Corr. Coeff. is Zero*

Theorem. The p.d.f. of the sample correlation coefficient r in a random sample of size n from an uncorrelated BVN population is given by

$$f(r) = \{B[\tfrac{1}{2}, \tfrac{1}{2}(n-2)]\}^{-1} (1 - r^2)^{(n-4)/2}, \ -1 \leq r \leq 1.$$

Proof. (Sawkin's method). Let $(X_1, Y_1), (X_2, Y_2), ..., (X_n, Y_n)$ be a random sample of size n drawn from uncorrelated BVN population with specifications.

$$E(X) = E(Y) = 0, \ \text{Var}(X) = \sigma^2, \ \text{Var}(Y) = \sigma'^2, \ \rho = 0.$$

We transform the random vector $\boldsymbol{Y}$ to the random vector $\boldsymbol{Z}$ by means of a linear orthogonal transformation, viz. $\boldsymbol{Z} = C\boldsymbol{Y}$, where C is an orthogonal matrix. Explicitly this means :

$$\left.\begin{array}{l} Z_1 = C_{11}Y_1 + C_{12}Y_2 + ... + C_{1n}Y_n \\ Z_2 = C_{21}Y_1 + C_{22}Y_2 + ... + C_{2n}Y_n \\ .... \ .... \ .... \ .... \ .... \ .... \ .... \\ Z_n = C_{n1}Y_1 + C_{n2}Y_2 + ... + C_{nn}Y_n \end{array}\right\} \text{ i.e. } \begin{bmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_n \end{bmatrix} = \begin{bmatrix} C_{11} & .... & C_{1n} \\ C_{21} & .... & C_{2n} \\ \vdots & \vdots & \vdots \\ C_{n1} & .... & C_{nn} \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}.$$

Let us choose $C_{11} = C_{12} = ... = C_{1n} = 1/\sqrt{n}$, whence Z_1 is given by

$$Z_1 = (Y_1 + Y_2 + ... + Y_n)/\sqrt{n} = \bar{Y}\sqrt{n} \Rightarrow Z_1^2 = n\bar{Y}^2.$$

Also, $\Sigma Z_i^2 = \Sigma Y_i^2 = \Sigma(Y_i - \bar{Y}) + \bar{Y}]^2 = \Sigma(Y_i - \bar{Y})^2 + n\bar{Y}^2$, $(1 \leq i \leq n)$ since $\Sigma(Y_i - \bar{Y}) = 0$.

Letting $\Sigma(X_i - \bar{X})^2 = nS_1^2$, $\Sigma(Y_i - \bar{Y})^2 = nS_2^2$, the preceding equation can be expressed as

$$\Sigma Z_j^2 = nS_2^2, \; j = 2, 3, \ldots, n. \qquad \ldots(i)$$

Now $r = \dfrac{\Sigma(X_i - \bar{X})(Y_i - \bar{Y})}{nS_1 S_2} = \dfrac{\Sigma(X_i - \bar{X})Y_i - \bar{Y}\,\Sigma(X_i - \bar{X})}{nS_1 S_2} = \dfrac{\Sigma(X_i - \bar{X})Y_i}{nS_1 S_2}$ (Def.) $[\Sigma(X_i - \bar{X}) = 0]$

If we choose $C_{2i} = (X_i - \bar{X})/S_1\sqrt{n}$, the above result gives

$$Z_2 = \Sigma C_{2i} Y_i = \Sigma(X_i - \bar{X})Y_i / S_1\sqrt{n} = rS_2\sqrt{n} \Rightarrow Z_2^2 = nS_2^2 r^2 \qquad \ldots(ii)$$

Eliminating Z_2^2 between (i) and (ii) we get

$$\Sigma Z_k^2 = n(1 - r^2)S_2^2, \; k = 3, 4, \ldots, n. \qquad \ldots(iii)$$

By Fisher's Lemma, $Z_i \sim N(0, \sigma'^2)$, $1 \le i \le n$, are independent and so $Z_i/\sigma' \sim N(0, 1) \Rightarrow (Z_i^2/\sigma'^2) \sim \chi^2_{(1)}$. Dividing (ii) and (iii) by σ'^2, and using Reproductive property of χ^2-variates.

$$U = (nS_2^2 r^2/\sigma'^2) \sim \chi^2_{(1)} \; ; V = n(1 - r^2)S_2^2/\sigma'^2 \sim \chi^2_{(n-2)}.$$

$$\therefore \qquad r^2 = [U/(U + V)] \sim \mathrm{B_I}[\tfrac{1}{2}, \tfrac{1}{2}(n - 2)]. \qquad [\S\ 21\text{-}61]$$

It follows that the elemental p.d.f. of r^2 is

$$dP(r^2) = \{B[\tfrac{1}{2}, \tfrac{1}{2}(n-2)]\}^{-1}(r^2)^{\frac{1}{2}-1}(1 - r^2)^{[(n-2)/2]-1}\,d(r^2), \; 0 \le r^2 \le 1$$

i.e. $$dP(r) = \{B[\tfrac{1}{2}, \tfrac{1}{2}(n-2)]\}^{-1}(1 - r^2)^{(n-4)/2}\,dr, \quad -1 \le r \le 1,$$

the factor 2 disappearing in converting the range $0 \le r^2 \le 1$ to $-1 \le r \le 1$.

Example : For a random sample of size 5 from an uncorrelated BVN population, if $P\{\,|r| \ge k\} = \alpha$, show that k is a root of the Eq.

$$k\sqrt{1 - k^2} + \sin^{-1}(k) - \tfrac{1}{2}\pi(1 - \alpha) = 0.$$

Solution. The p.d.f. of "r" from a BNV distribution ($\rho = 0$) is

$$f(r) = \{B[\tfrac{1}{2}, \tfrac{1}{2}(n-2)]\}^{-1}(1 - r^2)^{(n-4)/2}\,dr, \; -1 \le r \le 1. \qquad \ldots(i)$$

When $n = 5$, $B\,[1/2, 3/2] = \Gamma(1/2)\Gamma(3/2)/\Gamma(2) = \tfrac{1}{2}\pi \qquad \ldots(ii)$

$$\alpha = P\{|r| \ge k) = 1 - P\{|r| < k\} = 1 - P\{-k < r < k\} = 1 - 2P\{0 \le r \le k\}. \qquad \ldots(iii)$$

Because $f(r)$ is symmetric about $r = 0$. From (i), (ii) and (iii) we get

$$\frac{1}{2}(1 - \alpha) = P(0 \le r < k) = \frac{2}{\pi}\int_0^k (1 - r^2)^{1/2}\,dr = \frac{1}{\pi}\Big[r(1 - r^2)^{1/2} + \sin^{-1}(r)\Big]_0^k$$

or $$\tfrac{1}{2}\pi(1 - \alpha) = k\sqrt{1 - k^2} + \sin^{-1}(k) \Rightarrow k\sqrt{1 - k^2} + \sin^{-1}(k) - \tfrac{1}{2}\pi(1 - \alpha) = 0.$$

21-80. *Pearson's Chi-square Statistic*

In a large sample : $\displaystyle\sum_{j=1}^{k} \frac{(n_j - np_j)^2}{np_j} \sim \chi^2_{(k-1)}$ $\left[\text{Popularly written : } \displaystyle\sum_{i=1}^{n} \frac{(o_i - e_i)^2}{e_i} \sim \chi^2_{(k-1)}\right]$

where n_j is the observed frequency and np_j is the expected frequency of the jth class, C_j (say).

Proof. Consider a random sample of size n. The members of this sample are distributed at random in k classes $C_1, C_2, \ldots, C_k$. Let $p_j = \{\text{a sample observation} \in C_j\}$.

If n_j is the frequency of C_j, then these frequencies shall possess multinomial frequency function $f(n_1, n_2, \ldots, n_k) = f$, say ; where $f = [n!/n_1!n_2!\ldots n_k!)\, p_1^{n_1} p_2^{n_2} \ldots p_k^{n_k}$.

Here, $$\Sigma n_j = n; \quad \Sigma p_j = 1, 1 \le j \le k.$$

If the sample size is sufficiently large, n_j are large too and hence the Stirling's approximation for each of these factorials is valid. Then

$$f = \frac{(\sqrt{2\pi})\, e^{-n}\, n^{n+1/2}}{(\sqrt{2\pi})^k \,.\, e^{-(n_1+n_2+\ldots+n_k)}} \cdot \frac{p_1^{n_1}\, p_2^{n_2} \ldots p_k^{n_k}}{n_1^{n_1+\frac{1}{2}} \,.\, n_2^{n_2+\frac{1}{2}} \ldots n_k^{n_k+\frac{1}{2}}}$$

$$= \frac{(np_1/n_1)^{n_1+\frac{1}{2}} \ldots (np_k/n_k)^{n_k+\frac{1}{2}}}{(\sqrt{2\pi})^{k-1}\, (p_1\, p_i \ldots p_k)^{1/2}\, n^{(k-1)/2}} = A \prod_{i=1}^{k} \left(\frac{np_i}{n_i}\right)^{n_i+\frac{1}{2}} \qquad \ldots(1)$$

where $A^{-1} = (2n\pi)^{(k-1)/2} (p_1\, p_2 \ldots p_k)^{-1/2}$. We notice that A is independent of n_i's and for a given n and for a given set of theoretical probabilities p_i, A is constant. Further notice that $E(n_i) = np_i$. To simplify (1), we introduce

$$z_i = \frac{n_i - np_i}{\sqrt{np_i}} = \frac{n_i - \theta_i^2}{\theta_i}, \quad (\theta_i = \sqrt{np_i})\,. \qquad \ldots(2)$$

This gives : $n_i = (\theta_i^2 + \theta_i z_i)$, and substituting for n_i into (1) we get

$$\ln\left(\frac{f}{A}\right) \doteq \sum_{i=1}^{k} \left(n_i + \frac{1}{2}\right) \ln \frac{np_i}{n_i} = -\sum_{i=1}^{k} \left(\theta_i^2 + \theta_i z_i + \frac{1}{2}\right) \ln\left(1 + \frac{z_i}{\theta_i}\right).$$

Assuming that $(z_i/\theta_i) < 1$, we expand the logarithm to get

$$\ln \frac{f}{A} \doteq -\sum_{i=1}^{k} \left(\theta_i^2 + \theta_i\, z_i + \frac{1}{2}\right) \left\{\frac{z_i}{\theta_i} - \frac{1}{2}\frac{z_i^2}{\theta_i^2} + o\left(\frac{1}{\theta_i^3}\right)\right\} \doteq -\Sigma\, \{z_i \theta_i - \tfrac{1}{2} z_i^2 + z_i^2 + o\,(\theta_i^{-1})\} \qquad \ldots(3)$$

neglecting, of course, higher powers of (z_i/θ_i) if $z_i << \theta_i$. Now

$$\Sigma z_i \theta_i = \Sigma(n_i - np_i) = \Sigma n_i - n \Sigma p_i = n - n = 0. \qquad \ldots(4)$$

For large n, $O(\theta_i^{-1}) \to 0$; so Eq. (3) provides

$$\ln(f/A) \to -\tfrac{1}{2}\Sigma z_i^2, \text{ i.e. } f = A \exp[-(\Sigma z_i^2)/2] \qquad 1 \le i \le k. \qquad \ldots(5)$$

By hypothesis of multinomial frequency function, Z_i are independent variates and form (5) suggests that each Z_i is $N(0, 1)$, since A can be chosen so that $\int_0^1 df = 1$. In fact $A = (2\pi)^{-k/2*}$.

It follows that Σz_j^2 must be a $\chi^2_{(k-1)}$ variate, one degree of freedom having been lost by the linear constraint $\Sigma z_i \theta_i = 0$, by (4).

* Out objective is to find the probability that the random quantity χ^2 will exceed an arbitrary assigned number λ (say). As such the distribution of χ^2 is the total probability of all points inside the hypersphere $\Sigma z_i^2 = \lambda$. Consequently, we assume (5) as the *continuous* distribution and hence evaluate A by integration.

21-90. The Non-central Chi-square Distribution

Definition. Let $X_j \sim N(\mu_j, \sigma_j^2)$, $1 \le j \le n$ be independent variates. The r.v. $Y = (\Sigma X_j^2 / \sigma_j^2)$ is called non-central $\chi^2_{(n)}$; denoted by $\chi^2(n, \lambda)$ where $2\lambda = \Sigma u_i^2 / \sigma_i^2$.

To determine the p.d.f. of non-central χ^2 :

***Step* 1.** Let $X \sim N(\mu, \sigma^2)$. Then m.g.f. of $V = X^2/\sigma^2$ is [Example 16-30]

$$M_1(t) = (1-2t)^{-1/2} \exp\{2\lambda_0 t/(1-2t)\} = (1-2t)^{-1/2} e^{\lambda_0} \exp[\lambda_0/(1-2t)]. \qquad [\lambda_0 = \mu^2/2\sigma^2]$$

***Step* 2.** We now determine to m.g.f. of $Y = \Sigma(X_i^2/\sigma_i^2)$ and its p.d.f.

$$M(t:Y) = M\left(t : \Sigma \frac{X_i^2}{\sigma_i^2}\right) = \prod_{i=1}^{n} M\left(t : \frac{X_i^2}{\sigma_i^2}\right) = (1-2t)^{-n/2} \prod_{i=1}^{n} e^{-\lambda_i} \exp\frac{\lambda_i}{(1-2t)} \qquad \left[2\lambda_i = \frac{u_i^2}{\sigma_i^2}\right]$$

$$= (1-2t)^{-n/2} . e^{-\lambda} \exp[\lambda/(1-2t)] = (1-2t)^{-n/2} \exp[2\lambda t/(1-2t)]. \; [2\lambda = \Sigma(\mu_i^2/\sigma_i^2)]$$

where λ is called non-centrality parameter. Some authors write $\sigma = 2\lambda$.

Density. We use Inversion to Ch. Function for finding p.d.f. Thus

$$\phi_Y(t) = (1-2it)^{-n/2} \sum_{r=0}^{\infty} \frac{e^{-\lambda}}{r!}\left(\frac{\lambda}{1-2it}\right)^r = \sum_{r=0}^{\infty} \frac{e^{-\lambda}\lambda^r}{r!}(1-2it)^{-(r+n/2)} = \sum_{r=0}^{\infty} p_r (1-2it)^{-(r+n/2)}.$$

where $p_k = e^{-\lambda}\lambda^r / r!$.

$$\therefore \; f(y) = \frac{1}{2\pi}\int_0^{\infty} \phi_Y(t) e^{-ity} dt = \frac{1}{2\pi}\sum_{r=0}^{\infty} p_r \left\{\int_0^{\infty} e^{-ity}(1-2it)^{-(r+n/2)}\, dy\right\}. \text{ [Inversion Formula,}$$

$$\text{Recall : } Z \sim X^2_{n+2r}, \; f_Z(z) = \frac{e^{-z/2} . z^{(n/2+r)-1}}{\Gamma[(n+2r)/2].\, 2^{(n+2r)/2}}, \; \phi_z(t) = (1-2it)^{-(n+2r)/2}.$$

$$\therefore \; f(y) = \sum_{r=0}^{\infty} p_r \cdot \frac{e^{-y/2}\, y^{n/2+r-1}}{2^{r+(n/2)}\Gamma(r+n/2)} = e^{-\lambda}\frac{e^{-y/2}\, y^{n/2-1}}{2^{n/1}\Gamma(n/2)} + \sum_{r=1}^{\infty} p_r \frac{e^{-y/2}\, y^{n/2+r-1}}{2^{r+(n/2)}\Gamma[r+(n/2)]}.$$

Cumultants and moments :

$$\Sigma k_r t^r / r! = K(t:Y) = \ln M(t:Y) = \ln\{(1-2t)^{-n/2} . e^{-\lambda} . \exp\lambda/(1-2t)\}$$

$$= -(n/2)\ln(1-2t) - \lambda + \lambda(1-2t)^{-1}$$

$$= \left(\frac{n}{2}\right)\sum_{r=1}^{\infty}\frac{(2t)^r}{r} - \lambda + \lambda\sum_{r=0}^{\infty}(2t)^r = \frac{n}{2}\sum_{r=1}^{\infty} 2^r (r-1)!\frac{t^r}{r!} + \lambda\sum_{r=1}^{\infty} 2^r . r!\left(\frac{t^r}{r!}\right)$$

$$\therefore \quad k_r = 2^{r-1}(r-1)!(n+2\lambda r); \; k_{r+1} = \{2r[n+2(r+1)\lambda]/(n+2r\lambda)\}\, k_r \qquad \ldots(1)$$

Using logarithmic differentiation w.r.t. λ we get

$$dk_r/d\lambda = 2rk_r/(n+2\lambda r) \qquad \ldots(2)$$

From (1) & (2) : $k_{r+1} = [n+2(r+1)\lambda](dk_r/d\lambda)$

From 1(a) : $k_1 = n + 2\lambda \quad k_2 = 2(n+4\lambda) = \text{Variance}.$

Reproductive Property. If $X_i \sim \chi^2(\lambda_i, n_i)$, $i = 1, 2$ are independent variates, then

$$X_1 + X_2 \sim \chi^2(\lambda_1 + \lambda_2, n_1 + n_2).$$

Proof. $M(t : X_1 + X_2) = M(t : X_1)\, M(t : X_2)$

$$= (1-2t)^{-n_1/2} . e^{2\lambda_1 t/1-2t)} . (1-2t)^{-n_2/2}\, e^{-2\lambda_2 t/(1-2t)} = (1-2t)^{\frac{1}{2}(n_1+n_2)} . e^{2(\lambda_1+\lambda_2)t/(1-2t)}$$

This proves the property as stated.

Note. If $\lambda = 0$, then $M(t : \Sigma Y_i^2) = (1-2t)^{-n/2}$, so $f(u)$ is simply the χ^2-density function with n degrees of freedom, whence λ is called *non-centrality parameter*. A non-central chi-square variate with parameters λ and n is denoted by $\chi^2(\lambda, n)$.

21-91. *Miscellaneous Worked-out Problems*

Example 1. If X is $\chi^2_{(n)}$, show that for large n

(i) $(X-n)/\sqrt{2n} \sim N(0,1)$, (ii) $(\sqrt{2X} - \sqrt{2n}) \sim N(0,1)$.

Solution. Let $Y = (X-n)/\sqrt{2n}$.

$$M(t : Y) = M\,X/\sqrt{2n} - \sqrt{\tfrac{1}{2}n} = e^{-\sqrt{n/2}\,t}\, M(t/\sqrt{2n} : X) = e^{-t\sqrt{n/2}}\, [1 - 2t/\sqrt{2n}]^{-n/2}$$

$$\ln M(t : Y) = -\sqrt{n/2}\,t - \tfrac{1}{2} n \ln[1 - 2t/\sqrt{2n}]$$

$$= -\sqrt{\tfrac{1}{2}nt} + \frac{n}{2}\left[\frac{2t}{\sqrt{2n}} + \frac{4t^2}{2n}\cdot\frac{1}{2} + \frac{8t^3}{(\sqrt{2n})^3}\,\frac{1}{3} + \ldots\right] = \frac{t^2}{2} + \frac{\sqrt{2}}{3}\,\frac{t^3}{\sqrt{n}} + 0\left(\frac{1}{\sqrt{n}}\right).$$

As $n \to \infty$, $\ln M(t : Y) \to \frac{1}{2}t^2$. Since logarithm is a continuous function this provides

$$\lim M(t : Y) = e^{\frac{1}{2}t^2} \Rightarrow Y \sim N(0,1) \text{ for large } n.$$

(ii) $$P\{\sqrt{2X} - \sqrt{2n} < k\} = P\{\sqrt{2X} < k + \sqrt{2n}\} = P\{2X < k^2 + 2n + 2k\sqrt{2n}\}$$

$$= P\left\{\frac{X-n}{\sqrt{2n}} < k + \frac{k^2}{2\sqrt{2n}}\right\} \doteq P\{Y < k\}, \text{ for large } n.$$

Since $Y \to N(0, 1)$ as $n \to \infty$, $\sqrt{2X} - \sqrt{2n} \to N(0,1)$ as n is large.

Example 2. Let $X \sim \chi^2_{(n)}$ and define $U = (X-n)/\sqrt{2n}$. Using C.L.T. show that the limiting distributions of U is $N(0, 1)$.

Solution. (i) Let $X = X_1 + X_2 + \ldots + X_n$, where $X_{(i)} \sim \chi^2_{(1)}$ are i.i.d. variates. Then

$$\mathrm{E}(X) = n\mathrm{E}(X_i) = n; \quad \mathrm{Var}(X) = n\,\mathrm{Var}(X_1) = 2n. \text{ So}$$

$$U = S_n^* = \frac{S_n - \mathrm{E}(S_n)}{\sigma(S_n)} = \frac{X-n}{\sqrt{2n}} \to N(0,1) \text{ as } n \to \infty. \qquad \text{(by C.L.T.)}$$

Problems with Solutions Provided at the End of the Text

1*. Give a reasonable definition of χ^2-variate with zero degrees of freedom.

2*. Let $Y = \ln X$, where $X \sim \chi^2_{(n)}$. Show that $M(t : Y) = 2^t\Gamma(\frac{1}{2}n + t)/\Gamma(\frac{1}{2}n)$. If X_1 and X_2 are i.i.d. $\chi^2_{(n)}$-variates and $F = X_1/X_2$, deduce that for a positive integer k,

$$E(F^k) = \Gamma(\tfrac{1}{2}n + k)\Gamma(\tfrac{1}{2}n - k)/[\Gamma(\tfrac{1}{2}n)]^2.$$

3*. If $X_1, X_2, ..., X_n$ is a random sample from $N(\mu, \sigma^2)$, find the mean and variance of S^2.

Exercises

1. If X_1, X_2, X_3 is a random sample from $N(0, \sigma^2)$ population, find $P[\Sigma X_i^2 \leq 3\sigma^2\}$.
2. If $X_k \sim N(k, k^2)$, $k = 1, 2, 3$ are indep., construct a $\chi^2_{(3)}$, variate using only X_1, X_2, X_3.
3. Let X and Y denote the number of successes and failures respectively in n independent Bernoulli trials with success probability p. Show that
$$[(X - np)^2/np] + [(Y - nq)^2/nq]$$
can be approximated by $\chi^2_{(1)}$ for large n.
4. Give an example of two independent variates none of which is a χ^2-variate, although their sum is a χ^2-variate.
5. (a) If S^2 is the sample variance of a random sample of size 6 from the $N(\mu, 12)$, find $P\{2.3 < S_2 < 22.2\}$.
(b) For a random sample of size 25 from $N(3, 10^2)$ population, find
$$P\{0 < \overline{X} < 6, \quad 55.2 < S^2 < 145.6\}.$$
6. If $X_i \sim N(0, 1)$, $i = 1, 2$ are i.i.d. and $Z_j \sim N(1, 1)$, $j = 1, 2$ are i.i.d. variates, find the p.d.f. of
$$\tfrac{1}{2}[(X_1 - X_2)^2 + (Z_1 - Z_2)^2 + (Z_1 + Z_2)^2].$$
7. A point (X, Y, Z) is chosen at random in 3-dim space where X, Y, Z are i.i.d. $N(0, 1)$ variates. Find the prob. that the distance from the origin to the point will be < 1 unit.
8. If $X \sim$ Pois (λ) and $Y \sim \chi^2_{(2n)}$, show that for all positive integers n
$$P\{X \leq n - 1\} = P\{Y \geq 2\lambda\}.$$
9. Prove that $P\{\chi^2_{(2n+2)} > 2\lambda\} = \dfrac{1}{n!}\int_\lambda^\infty e^{-y}\, y^n\, dy = \sum_{r=0}^{n} \dfrac{e^{-\lambda}\lambda^r}{r!}$, where $y = \frac{1}{2}\chi^2$. Explain the uses of this result.
10. It $X \sim \chi^2_n$, show that $1 - P(X \leq x) \leq n/x$.
11. If $X_1, X_2, ..., X_n$ are i.i.d. Expo (λ) variates, show that $2\lambda\, \Sigma X_i \sim \chi^2_{(2n)}$.
12. If $X \sim \chi^2_{(n)}$, find the p.d.f. of $Y = X/(1 + X)$.
13. If $X_1, X_2, ..., X_n$ are independent variates with continuous c.d.f. $F_1, F_2, ..., F_n$. Show that
$$-2\ln[F_1(X_2)\, F_2(X_i) \ldots F_n(X_n)] \sim \chi^2_{(2n)}.$$

14. Let $f(x, n)$ and $F(x, n)$ denote the p.d.f. and the c.d.f. of $\chi^2_{(n)}$ variate respectively and $\Phi(x)$ be c.d.f. of $N(0, 1)$ variate. Show that

(a) $xf(x, n) = nf(x\,;n+2)$. (b) $\Sigma f(x, 2n) = 1/2$, $(n = 1, \ldots \infty)$.

(c) $F(x, n) - F(x\,;n+2) = (\frac{1}{2}x)^{n/2}\, e^{-x/2} / \Gamma(\frac{1}{2}n+1)$.

(d) $\Sigma f(x, 2n+1) = \Phi(\sqrt{2x}) - 1/2,\ x > 0$.

15. If $X \sim$ Pois (λ) and $2\alpha\lambda \sim \chi^2_{(2n)}$, obtain the unconditional distribution of X. Give the name of this distribution and find its mean.

16. Let X, Y, Z be χ^2-variates with $2a, 2b, 2c$, d.f. respectively.

(i) Show that $X/(X+Y)$ is independently distributed of $(X+Y)/(X+Y+Z)$.

(ii) Obtain the joint p.d.f. of the distribution of $U = X/(X+Y+Z)$ and $V = X/(X+Y+Z)$. Name the distribution and find Corr (U, V).

17. If X and Y are i.i.d. $\chi^2_{(2)}$-variates, show that p.d.f. of $Z = \frac{1}{2}(X - Y)$ is

$$f(z) = \frac{1}{2} e^{-|z|}, \quad -\infty < z < \infty.$$

18. Let $X \sim N(0, 1)$ and $Y \sim \chi^2_{(1)}$ be independent. Show that

$$(aX + bY)^2 / (a^2 + b^2) \sim \chi^2_{(1)}, \quad (a \neq 0,\ b \neq 0).$$

19. If $X \sim N(0, 1)$ and $Y \sim \chi^2_{(1)}$ are independent, show that $W = XY^{1/2}$ is distributed like the sum of n indep. variates W_j with p.d.f.

$$f(w_j) = \frac{1}{2}\exp(-|w_j|), \quad -\infty < w_j < \infty.$$

20. Let $X_1, X_2, \ldots, X_n$ be i.i.d. $N(0, 1)$ variates. If $X = \Sigma a_i X_i$ and $Y = \Sigma X_i^2$ $(i = 1, 2, \ldots, n)$, show that the conditional distribution of Y given that $X = 0$ is $\chi^2_{(n-1)}$ distribution.

21. Let $y\,dx$ be the probability that X lies between x and $x + dx$, and y be given by the solution of differential equation $(dy/dx) = y(a - x)/(bx + c)$.
With suitable values of the constants a, b, c show that certain linear function of X has the $\chi^2_{(1)}$-distribution where $n = 2[1 + (a/b) + (c/b^2)]$.

22. Find the limiting distribution of $n^{-2}\chi^2_{(n)}$.

23. If $X \sim \chi^2_{(n)}$, show that $\sqrt{2X} \sim N(\sqrt{(2n-1)}, 1)$ for large n.

24. If $X \sim B_1(n, \lambda)$ where n is an integer, show that for large n and small λ, a first approximation gives $-2\lambda \ln X$ as $\chi^2_{(2n)}$-distribution. Show further that as an improved approximation, $-(2\lambda + n - 1)\ln X$ is $\chi^2_{(2n)}$-distributed.

25. Let $X_1, X_2, \ldots, X_k$ be i.i.d. $\chi^2_{(m)}$-variates. Define $Y_1 = X_2/X_1$, $Y_2 = X_3/X_1, \ldots, Y_{k-1} = X_k/X_1$. Find the joint p.d.f. of $Y_1, Y_2, \ldots, Y_{k-1}$.
26. Let $X_1, X_2, \ldots, X_n$ be a random sample from bin $(1, p)$, $0 < p < 1$.

    Find the p.d.f. of the sample variance $\hat{S}^2$.
27. Let $X_1, X_2, \ldots, X_n$ be a random sample from Pois (λ). Find Var $(\overline{X})$ and Var $(\hat{S}^2)$.
28. Let $X_1, X_2, \ldots, X_n$ be a random sample from $N(\mu, \sigma^2)$. Find Var $(\hat{S}^2)$.
29. Let $X_1, X_2, \ldots, X_n$ be a random sample from $N(0, 1)$. Define

$$\overline{X}_k = \frac{1}{k} \sum_{i=1}^{k} X_i,\ \overline{X}_{n-k} = \frac{1}{n-k} \sum_{k+1}^{n} X_i,\ \hat{S}_k^2 = \frac{1}{k-1} \sum_{i=1}^{k} (X_i - \overline{X}_k)^2,\ [\hat{S}_{n-k}]^2 = \frac{1}{(n-k-1)} \sum_{k+1}^{n} (X_i - \overline{X}_{n-k})^2.$$

    Find the p.d.f. of (a) $k\overline{X}_k^2 + (n-k)\,\overline{X}_{n-k}^2$, (b) $(k-1)\,\hat{S}_k^2 + (n-k-1)\,\hat{S}_{n-k-1}^2$.
30. Let $X_1, X_2, \ldots, X_m$ and $Y_1, Y_2, \ldots, Y_n$ be indep. random samples from $N(0, \sigma^2)$ population, their means are $\overline{X}, \overline{Y}$ and variances $\hat{S}_x^2, \hat{S}_y^2$ respectively. Let the pooled variance $\hat{S}_p^2$ be defined by $\hat{S}^2 = [(m-1)\,\hat{S}_x^2 + (n-1)\,\hat{S}_y^2]/(m+n-2)$. Prove that

    $(\overline{X}, \overline{Y}) \sim N\{0, \sigma^2[(1/m) + (1/n)]\}$ and $(m+n-2)\,\hat{S}^2/\sigma^2 \sim \chi^2_{(m+n-2)}$, and that they are indep.
31. Prove that each linear constraint on $[f_i]$ reduces by unity the number of degrees of freedom of the chi-square $\Sigma[(f_i - e_i)^2/e_i]$, where $e_i = \mathrm{E}(f_i)$.

Success always occurs in private, and failure in full view.

The secret of creativity is knowing how to hide your sources.

(Albert Einstein)

Fisher-Student t-Distribution. Snedecor-Fisher F-Distribution 22

Introduction. In this chapter, we consider the *basic concepts* of two outstanding sampling distributions called t-distribution and F-distribution which are related to Normal and Ch-square distributions. The wide variety of applications of these distributions may be considered later on.

FISHER-STUDENT t-DISTRIBUTION

22-10. *Definition*

A random variable T is said to possess t-distribution with n degrees of freedom, if its p.d.f. is given by

$$f(t) = \frac{1}{\sqrt{n}\, B\left(\frac{1}{2}n, \frac{1}{2}\right)} \left(1 + \frac{t^2}{n}\right)^{-(n+1)/2}, \quad -\infty < t < \infty.$$

Variate T having p.d.f. (1) is briefly written, $T \sim t_{(n)}$. See § 22-32 for sketch of $t_{(3)}$ and $t_{(17)}$.

22-11. *Fisher's t-Density*

Let $X \sim N(0, 1)$ and $Y \sim \chi^2_{(n)}$ be independent variates. Then, the variate T defined by

$$\boxed{T = \frac{X}{\sqrt{Y/n}} = \frac{N(0,1)}{\sqrt{\chi^2_{(n)}/n}}} \qquad \text{[Fisher's } t\text{-statistic]} \qquad ...(1)$$

has (Fisher's) t-distribution with n degrees of freedom.

Proof. Since X and Y are independent, their joint elemental p.d.f. is

$$dP_1(x, y) = f_1(x)\, dx\, f_2(y)\, dy = \frac{e^{-x^2/2}\, dx}{\sqrt{2\pi}} \frac{e^{-y/2}\, y^{\frac{1}{2}n-1}\, dy}{\Gamma(\frac{1}{2}n).(2)^{n/2}}, \quad -\infty < x < \infty,\ 0 < y < \infty$$

Put $y = s$, $t = x/(y/n)^{1/2}$, i.e. $x = t(s)^{1/2}/\sqrt{n}$, $y = s$; $[s > 0, |t| < \infty]$

$$\therefore \qquad |J| = \frac{\partial(x, y)}{\partial(s, t)} = \begin{vmatrix} \partial x/\partial s & \partial x/\partial t \\ \partial y/\partial s & \partial y/\partial t \end{vmatrix} = \begin{vmatrix} t/2\sqrt{ns} & (s/n)^{1/2} \\ 1 & 0 \end{vmatrix} = \left(\frac{s}{n}\right)^{1/2}$$

The Joint p.d.f. in terms of s and t is thus given by

$$dP_2(s, t) = \frac{e^{-(st^2/2n)}}{\sqrt{2\pi}} \frac{e^{-\frac{1}{2}s}\, s^{-\frac{1}{2}n-1}}{\Gamma(n/2)\,(2)^{n/2}} \left(\frac{s}{n}\right)^{1/2} ds\, dt = \frac{s^{\frac{1}{2}(n+1)-1}\, e^{-\frac{1}{2}[1+(t^2/n)]s}\, ds\, dt}{\sqrt{2\pi}\,\Gamma(n/2)\,(2)^{n/2}}.$$

To obtain the individual p.d.f. of T, we integrate **out** unwanted s to obtain

$$dP(t) = dt \int_0^\infty \frac{e^{-\frac{1}{2}s[1+t^2/n]}\, s^{\frac{1}{2}(n+1)-1}\, ds}{(2n\pi)^{1/2}\,\Gamma(\frac{1}{2}n)\,(2)^{n/2}} \qquad \text{[Use working gamma integral]}$$

$$= \frac{\Gamma[\frac{1}{2}(n+1)]\, dt}{\left\{\frac{1}{2}[1+(t^2/n)]\right\}^{(n+1)/2} (2n\pi)^{1/2}\,\Gamma(\frac{1}{2}n)\,(2)^{n/2}} = \frac{\Gamma[\frac{1}{2}(n+1)]}{\sqrt{\pi}\,\Gamma(\frac{1}{2}n)} \frac{1}{\sqrt{n}} \frac{dt}{[1+(t^2/n)]^{(n+1)/2}}\,; |t| < \infty$$

Thus, the p.d.f. of T is given by

$$f(t) = [\sqrt{n}\, B(\tfrac{1}{2}n, \tfrac{1}{2})]^{-1}\, [1+t^2/n]^{-(n+1)/2}\,; \quad -\infty < t < \infty. \qquad \text{...(2)}$$

22-12. *William Gosset's Student Law*

If $X_1, X_2, \ldots, X_n$ is a random sample from $N(\mu, \sigma^2)$ population, then

$$T = [(\overline{X} - \mu)\sqrt{n}/\hat{S}] \sim t_{(n-1)}. \qquad \text{[Student's } t\text{-statistic]} \qquad \text{...(1)}$$

Proof. $Y = \left(\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}}\right)$ is $N(0, 1)$; $Z = \dfrac{(n-1)\,\hat{S}^2}{\sigma^2}$ is $\chi^2_{(n-1)}$, Y and Z are independent; hence

$$T = Y/[Z/(n-1)]^{1/2} = (\overline{X} - \mu)/(\hat{S}/\sqrt{n})$$

Thus, t is $t_{(n-1)}$ distributed [by § 22-11 above]

Aliter. We rewrite (1) as

$$\frac{T^2}{n-1} = \left(\frac{(X-\mu)^2}{(\sigma^2/n)}\right) \Big/ \left[\frac{(n-1)\,\hat{S}^2}{\sigma^2}\right] = \frac{Y^2}{Z}.$$

Now $Y = \dfrac{(X-\mu)}{\sigma/\sqrt{n}}$ is $N(0, 1)$, hence $Y^2 = \dfrac{(X-\mu)^2}{\sigma^2/n}$ is $\chi^2_{(1)}$. Also, (nS^2/σ^2) is $\chi^2_{(n-1)}$.

Further $\overline{X}$ and S^2 are independent distributed [§ 21-70], it follows that the above $T^2/(n-1)$ is the ratio of two independent, $\chi^2_{(1)}$ and $\chi^2_{(n-1)}$ variates, and as such $T^2/(n-1)$ is $B_{II}[(\frac{1}{2}, \frac{1}{2}(n-1)]$-variate [§ 21-60]. Consequently, its probability law is given by

$$dF(t) = [B(\tfrac{1}{2}, \tfrac{1}{2}(n-1)]^{-1} \left(1 + \frac{t^2}{n-1}\right)^{-(n-1+1)/2} \left(\frac{t^2}{n-1}\right)^{(1/2)-1} d\left(\frac{t^2}{n-1}\right)$$

$$= \frac{1}{B[\frac{1}{2}, \frac{1}{2}(n-1)]} \frac{1}{(n-1)^{1/2}} \left(1 + \frac{t^2}{n-1}\right)^{-n/2} dt\,; \quad -\infty < t < \infty. \qquad \text{...(2)}$$

We have dropped the factor 2 and taken the range $-\infty < t < \infty$, since $F(\infty) = 1$, $F(-\infty) = 0$.

Cor. For one-degree of freedom [$n = 1$] in definition or $n = 2$ in (2)]

$$f(t) = [B(\tfrac{1}{2}, \tfrac{1}{2})]^{-1} (1+t^2)^{-1} = 1/\pi(1+t^2), \quad -\infty < t < \infty.$$

It follows that $t_{(1)}$ is a standard Cauchy-variate.

Remarks. Fisher's T-statistic is defined by Eq. 22-11(1) and Student's T-statistic is defined by Eq. 22-12(1). They have the same p.d.f. except that for Fisher's t-distribution, d.f. is n and for Student's t-distribution, d.f. is $(n-1)$. We may note that Student's t-distribution may be regarded as a special case of Fisher's t-distribution.

Example : Let X, Y, Z, W be i.i.d. $N(0, \sigma^2)$ variates and define $T = k(X+Y)/\sqrt{Z^2+W^2}$. For what value of k, $T \sim t_{(n)}$. Find n as well.

Solution. $(X+Y) \sim N(0, 2\sigma^2)$, hence $U = (X+Y)/\sqrt{2}\sigma$ is $N(0, 1)$. $V = (Z^2+W^2)/\sigma^2 \sim \chi^2_{(2)}$.

$$\therefore \qquad T_1 = \frac{(X+Y)\sqrt{2}\sigma}{\sqrt{(Z^2+W^2)/2\sigma^2}} = \frac{N(0,1)}{\sqrt{\chi^2_{(n)}/n}} \sim t_{(2)}.$$

On comparison with T, we find $k = 1$, $n = 2$. And $T \sim t_{(2)}$.

22-13. *Relation of T with Normal Distribution*

As $n \to \infty$, $t_{(n)} \to N(0, 1)$.

Proof.
$$f(x) = k[1+(x^2/n)]^{-(n/1)/2}, \quad k^{-1} = \sqrt{n}\, B(\tfrac{1}{2}, \tfrac{1}{2}n) \qquad \text{...(1)}$$

$$\therefore \qquad \lim_{n\to\infty} k = \lim_{n\to\infty} \frac{1}{\sqrt{n}} \frac{\Gamma(\frac{1}{2}+\frac{1}{2}n)}{\Gamma(\frac{1}{2})\Gamma(\frac{1}{2}n)} = \frac{1}{\sqrt{2\pi}} \qquad \text{...(2)}$$

$$\lim_{n\to\infty}\left(1+\frac{x^2}{n}\right)^{-(n+1)/2} = e^{-x^2/2}. \left[\lim_{n\to\infty}\left(1+\frac{a}{n}\right)^{bn+c} = e^{ab}\right]. \qquad \text{...(3)}$$

Taking limits in (1), using the limiting results (2) and (3) we obtain

$$\lim f(x) = (\sqrt{2\pi})^{-1} \exp(x^2/2), \quad -\infty < x < \infty.$$

Hence for large n, $t_{(n)} \to N(0, 1)$.

Remark. The t-family ranges from the heavy-tailed Cauchy distribution to the Normal distribution. [See Fig. § 22-32.]

22-20. *Non-existence of m.g.f.*

If $X \sim t_{(n)}$, $(n > 1)$ it can be shown that $\mathrm{E}(\mid X \mid^r) < \infty$, for $r < n$ and $\mathrm{E}(\mid X \mid^r) = \infty$, for $r \geq n$. This means, the first $(n-1)$ moments of X exist but moments of order $r \geq n$ do not exist. Consequently, the m.g.f. of X does not exist.

22-21. *Moments of t-distribution*

If X is $N(0, 1)$ and Y is $\chi^2_{(n)}$, X and Y independent, then $T = X + (Y/n)^{1/2} = n^{1/2} X/(Y)^{1/2} \sim t_{(n)}$

$$\therefore \quad \mathrm{E}(T^r) = n^{r/2}\,\mathrm{E}(X^r)\,\mathrm{E}(Y^{-r/2}).$$

Since X is symmetric about zero, its odd moments are zero; hence only even moments of T are non-zero. So $(r \to 2r)$

$$\mathrm{E}(T^{2r}) = n^r\,\mathrm{E}(X^{2r})\,\mathrm{E}(Y^{-r}) = n^r\,\mathrm{E}(Z^r)\,\mathrm{E}(Y^{-r}) \qquad [Z = X^2 \sim \chi^2_{(1)}] \qquad \text{...(i)}$$

If $W \sim \chi^2_{(n)}$, then $\mathrm{E}(W^k) = 2^k\,\Gamma(\tfrac{1}{2}n + k)/\Gamma(\tfrac{1}{2}n) = 2^k(\tfrac{1}{2}n)^{[k]}$. [§ 21-20]

$$\therefore \quad \mathrm{E}(Y^{-r}) = 2^{-r}\,\Gamma(\tfrac{1}{2}n - r)/\Gamma(\tfrac{1}{2}n);\ \mathrm{E}(Z^r) = 2^r(\tfrac{1}{2} + r)/\Gamma(\tfrac{1}{2}).$$

Substituting into (i) we recover

$$\mu_{2r} = \mathrm{E}(T^{2r}) + \frac{n^r\,\Gamma(\tfrac{1}{2}n - r)\,\Gamma(\tfrac{1}{2} + r)}{\Gamma(\tfrac{1}{2})\,\Gamma(\tfrac{1}{2}n)} = \frac{n^r\,(\tfrac{1}{2})^{[r]}}{(\tfrac{1}{2}n - r)^{[r]}} = n^r\,\frac{1.3.5\ldots(2r-3)(2r-1)}{(n-2)(n-4)\ldots(n-2r)}. \qquad \text{...(1)}$$

Aliter. $f(x) = k[1 + (x^2/2n)]^{-(n+1)/2};\ - < x < \infty,\ [k^{-1} = \sqrt{n}\,B(\tfrac{1}{2}n, \tfrac{1}{2})]$

Since X is symmetric about zero, all its odd moments are zero ; in particular $\mathrm{E}(X) = 0$. Thus, the simple and central moments are identical. Now,

$$\mu_{2r} = k\int_{-\infty}^{\infty} x^{2r}\,[1 + (x^2/n)]^{-(n+1)/2}\,dx.$$

Let $x^2/n = z$, then $dx = \tfrac{1}{2}(nz^{-1})^{1/2}\,dz;\ -\infty < x < \infty \ \Rightarrow\ 0 < z < \infty.$

$$\therefore\ \mu_{2r} = \frac{2}{\sqrt{n}\,B(\tfrac{1}{2}n, \tfrac{1}{2})}\int_0^{\infty}(nz)^r\,(1+z)^{-(n+1)/2}\,\frac{\sqrt{n}\,dz}{2\sqrt{z}},$$

$$= \frac{n^r}{B(\tfrac{1}{2}n, \tfrac{1}{2})}\int_0^{\infty}\frac{z^{(r+\frac{1}{2})-1}\,dz}{(1+z)^{(n+1)/2}}, \qquad \left[\int_0^{\infty}\frac{x^{a-1}\,dx}{(1+x)^{(a+b)}} = B(a, b)\right]$$

$$= n^r\,B(r + \tfrac{1}{2}, \tfrac{1}{2}n - r)/B(\tfrac{1}{2}n, \tfrac{1}{2}) = n^r\,\Gamma(\tfrac{1}{2}n - r)\,\Gamma(r + \tfrac{1}{2})/\Gamma(\tfrac{1}{2}n)\,\Gamma(\tfrac{1}{2}). \qquad \text{...(1)}$$

$$= (n)^r\,.\,(\tfrac{1}{2})^{[r]}/(\tfrac{1}{2}n - k)^{[r]}. \qquad [a^{[r]} = a(a+1)\ldots(a+r-1)] \qquad \text{...(2)}$$

22-22. *Moment Recurrence Formula and Pearson's Coefficients*

Replacing r by $(r - 1)$ in 22-21 (1) we get

$$\mu_{2r-2} = n^{r-1}\,\Gamma(\tfrac{1}{2}n - r + 1)\,\Gamma(r - \tfrac{1}{2})/\Gamma(\tfrac{1}{2}n)\,\Gamma(\tfrac{1}{2})$$

Dividing the result of μ_{2r} by that of μ_{2r-2} we get

$$\mu_{2r} = [n(2r-1)/(n-2r)]\,\mu_{2r-2}. \qquad \text{...(1)}$$

In particular : $\quad \mathrm{Var}\,(T) = \mu_2 = n/(n-2), \ (n > 2)$

Since $\mu = 0$, the coefficient of variation for T does not exist.

Since $\mu_3 = 0$, so $\beta_1 = 0$, $\gamma_1 = 0$. Also $\mu_4 = 3n^2/(n-2)(n-4)$, $(n>4)$

So $\beta_2 = 3(n-2)/(n-4)$, $\gamma_2 = 6/(n-4)$.

Obviously, (β_1, β_2) lies on the St. line : $\beta_1 + \beta_2 = 3(n-2)/(n-4)$.

Note.
$$\lim_{n\to\infty} \beta_2 = \lim_{n\to\infty} \frac{3[1-(2/n)]}{[1-(4/n)]} = 3.$$

This shows that $t_{(n)}$ tends to Normal distribution as $n \to \infty$. In practice when n is large, say $n > 30$. $t_{(n)}$ is approximated by $N(0, 1)$.

22-30. *Mode and Points of Inflexion*

Here $y = f(x) = k\,[1+(x^2/n)]^{-(n+1)/2}$, $k^{-1} = n^{1/2} B(\tfrac{1}{2}n, \tfrac{1}{2})]$, $-\infty < x < \infty$

We take logarithm of both sides and then differentiate

$$\ln y = \ln k - \tfrac{1}{2}(n+1)\ln[1+(x^2/n)]$$

$$y'/y = -(n+1)\,x/(n+x^2);\ (y''/y) - (y'/y)^2 = -(n+1)(n-x^2)/(n+x^2)^2. \qquad \ldots(1)$$

For max-min of y, $y' = 0$ and 1(a) provides $x = 0$. From 1(b)

$$y''(0) = -[k\,.(n+1)/n] < 0.$$

Hence the modal value of $t_{(n)}$-variate is at $x = 0$.

For determining points of inflexion, we put $y'' = 0$ and eliminate (y'/y) from 1(a) and 1(b). Thus,

$$\frac{(n+1)^2 x^2}{(n+x^2)^2} = \frac{(n+1)(n-x^2)}{(n+x^2)^2} \Rightarrow x^2 = \frac{n}{n+2} \Rightarrow x = \pm\left[\frac{n}{(n+2)}\right]^{1/2}$$

So, the points of inflexion are given by $\left[\pm\left(\dfrac{n}{n+2}\right)^{1/2},\ k\left(\dfrac{n+2}{n+3}\right)^{(n+1)/2}\right].$

Remark. As $n \to \infty$, these points tend to $[\pm 1, (2\pi e)^{-1/2}]$ which are the points of inflexion for $N(0, 1)$.

22-31. *Mean Absolute Deviation from Mean*

Here : $$f(x) = k[1+(x^2/n)]^{-(n+1)/2},\ [k = n^{-1/2} B(\tfrac{1}{2}n, \tfrac{1}{2})],\ -\infty < x < \infty \qquad \ldots(1)$$

Since f is symmetric about $x = 0$, $[f(-x) = f(x)]$; $E(X) = 0$. Now

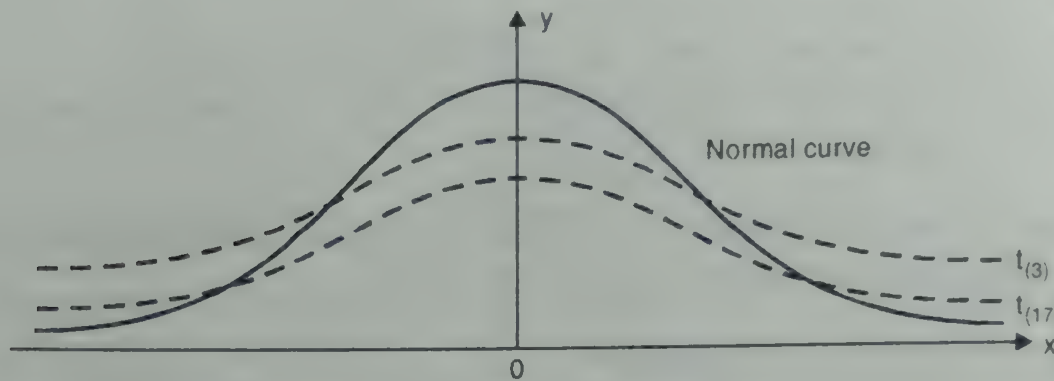
$$M = E(|X - 0|) = k\int_{-\infty}^{\infty} \frac{|x|\,dx}{[1+(x^2/n)]^{(n+1)/2}} = 2k\int_0^{\infty} \frac{x\,dx}{[1+(x^2/n)]^{(n+1)/2}}$$

To evaluate it, we put $x^2/n = z$; then $2x\,dx = n\,dz$ and so

$$M = nk\int_0^{\infty} \frac{dz}{(1+z)^{(2n+1)/2}} = \frac{2nk}{n-1} = \frac{\sqrt{n}\,\Gamma[\tfrac{1}{2}(n+1)]}{\Gamma(\tfrac{1}{2})\Gamma(\tfrac{1}{2}n)} \cdot \frac{2}{n-1} = \frac{\sqrt{n}\,\Gamma[\tfrac{1}{2}(n-1)]}{\sqrt{\pi}\,\Gamma(\tfrac{1}{2}n)}. \qquad \text{[Putting for } k]$$

22-32. *Shape of t-density Curve*

$$f(x) = k[1+(x^2/n)]^{-(n+1)/2}; \ -\infty < x < \infty$$



(1) The curve is symmetrical about $x = 0$, $[f(-x) = f(x)]$.

(2) As $x \to \pm\infty$, $f(x) \to 0$; thus x-axis in an asymptote to the curve $y = f(x)$.

(3) $\mu_2 = n/(n-2)$, $\beta_2 = 3(n-2)/(n-4)$.

For $n > 2$, $\mu_2 > 1$, i.e. Var (X) > Var $\{N(0, 1))\}$; for $n > 4$, $\beta_2 > 3 \Rightarrow t_n$ curve is flatter on the top than the $N(0, 1)$-curve. As $n \to \infty$, t_n-curve coincides with the Normal curve.

(4) The points of inflexion are : $\{\pm [n/(n+2)]^{1/2},\ k\,[(n+2)/(n+3)]^{(n+1)/2}\}$ and mode occurs at $x = 0$, at a height k above origin. A rough trace is shown above.

22-33. *Probability Tables for t-distribution*

Student (William Gosset) prepared tables for $P_S = P\{t \le t_0\}$ for different values of t_0 and different d.f. $(n = 1, 2, ...)$. Fisher and Yates prepared $P_F = P\{|\,t\,| > t_0\}$ for different value of t_0 and different d.f. $(n = 1, 2, ...)$. The relation between P_S and P_F, using symmetry of t-distribution is

$$P_S = P\{t \le t_0\} = P\{-\infty < t \le t_0\} = 0.5 + P\{0 < t \le t_0\} \qquad ...(1)$$

$$P_F = P\{|t| \ge t_0\} = 1 - P\{|t| < t_0\} = 1 - 2P\{0 < t \le t_0\} \qquad ...(2)$$

$$\therefore \quad 2P_S + P_F = 2 \Rightarrow P_F = 2(1 - P_S). \qquad ...(3)$$

The Appendix gives the values of t for different d.f.'s. These are used in the problems of Statistical Inference.

22-34. *Worked-out Problems*

***Example* 1.** Show that the mean value of the positive square root of a Beta variate of the second kind with parameters l and m is $\Gamma(l+\frac{1}{2})\,\Gamma(m-\frac{1}{2})/\Gamma(l)\,\Gamma(m)$.

Hence show that the mean value of $|\,t\,|$ for ν degrees of freedom is

$$\sqrt{\nu}\,\Gamma[\tfrac{1}{2}(\nu-1)]/\sqrt{\pi}\,\Gamma(\tfrac{1}{2}\nu).$$

Solution. If $X \sim B_{II}(l, m)$, then its probability density $f(x)$ is

$$f(x) = \frac{x^{l-1}}{B(l, m)\,(1+x)^{l+m}}; \quad x \ge 0,\ l,\ m > 0$$

$$\mathsf{E}(\sqrt{X}) = \int_0^\infty \frac{x^{1/2} . x^{l-1}\, dx}{B(l, m)(1+x)^{l+m}} = \frac{1}{B(l, m)} \int_0^\infty \frac{x^{(l+1/2)-1}\, dx}{(1+x)^{(l+1/2)+(m-1/2)}}$$

$$= \frac{B(l+\frac{1}{2}, m-\frac{1}{2})}{B(l, m)} = \frac{\Gamma(l+\frac{1}{2})\,\Gamma(m-\frac{1}{2})}{\Gamma(l) . \Gamma(m)}.$$

We recall that $(t^2/\nu) \sim \mathrm{B}_{\mathrm{II}}(\frac{1}{2}, \frac{1}{2}\nu)$; hence putting $(\nu X)^{1/2} = |t|, l = \frac{1}{2},\ m = \nu/2$ we get

$$\mathsf{E}(|t|) = \sqrt{\nu}\,\mathsf{E}(\sqrt{X}) = \sqrt{\nu}\,\Gamma(1)\,\Gamma[(\nu-1)/2]/\Gamma(\tfrac{1}{2})\,\Gamma(\tfrac{1}{2}\nu) = \sqrt{\nu}\,\Gamma[\tfrac{1}{2}(\nu-1)]/\sqrt{\pi}\,\Gamma(\nu/2).$$

***Example* 2.** Let the random variables X and Y be independent and follow $\chi^2_{(n)}$ distribution. Show that $U = \sqrt{n}\,(X-Y)/2\sqrt{XY}, V = X+Y$ are independent and recognise their distribution.

Solution. Let $u = \sqrt{n}\,(x-y)/2\sqrt{xy},\ v = x+y$. To ease evaluations, put $x = r\cos^2\theta$, $y = r\sin^2\theta$ so that $|J| = 2r\sin\theta\cos\theta$. Thus, $u = \sqrt{n}(\cos^2\theta - \sin^2\theta)/2\sin\theta\cos\theta$ $= \sqrt{n}\cot 2\theta, v = r$. Now

$$dF(x, y) = f_1(x) f_2(y)\, dx\, dy = (xy)^{(n/2)-1} e^{-(x+y)/2}\, dx\, dy / 2^n\, \Gamma^2(n/2)$$

$$dF(r, \theta) = \left\{\frac{r^{n-1} e^{-r/2}}{\Gamma(n)\, 2^n}\, dr\right\}\left\{\frac{2\Gamma(n)}{\Gamma^2(n/2)}\,\frac{(2\sin\theta\cos\theta)^{n-1}\, d\theta}{2}\right\}$$

Setting $\cot 2\theta = u/\sqrt{n}$, i.e. $\sin 2\theta = [1+u^2/n]^{-1/2}, d\theta = (2\sqrt{n})^{-1}(1+u^2/n)^{-1}\, du$

$$\therefore \qquad dF(r, u) = \left\{\frac{r^{n-1} e^{r/2}}{2^n\,\Gamma(n)}\, dr\right\}\left\{\frac{\Gamma(n)}{\Gamma^2(n/2)}\,\frac{1}{\sqrt{n}}\left(1+\frac{u^2}{n}\right)^{-(n+1)/2}\frac{du}{2^{n-1}}\right\}.$$

This shows that U and V (=R) are independent distributed. Obviously, $V \sim \chi^2_{(2n)}$, which we otherwise known as well (Reproductive Property). To find nature of U, we use Legendre. Duplication formula :

$$\sqrt{\pi}\,\Gamma(n) = 2^{n-1}\,\Gamma(n/2)\,\Gamma[(n+1)/2].$$

$$\therefore \qquad \Gamma(n)/\Gamma^2(n/2) = 2^{n-1}\,\Gamma(\tfrac{1}{2}+\tfrac{1}{2}n)/\Gamma(\tfrac{1}{2})\,\Gamma(n/2) = 2^{n-1}/B(\tfrac{1}{2}, \tfrac{1}{2}n).$$

$$\therefore \qquad f_U(u) = [1+(u^2/n)]^{-(n+1)/2}/\sqrt{n}\, B(\tfrac{1}{2}, \tfrac{1}{2}n), \ -\infty < u < \infty.$$

Thus, $U \sim t_{(n)}$. Note that

$$U = \sqrt{n}\,[\sqrt{X/Y} - \sqrt{Y/X}]/2 = \sqrt{n}\ [(F_{n,n})^{1/2} - (F_{n,n})^{-1/2}]/2.$$

Problems with Solutions Provided at the End of the Text

1*. Express the constants C, a, m of the following distribution in terms of its σ^2 and β_2

$$dF(x) = C[1-(x^2/a^2)]^m \, dx, \quad -a < x < a. \qquad \ldots(1)$$

Show that if X is related to a variable t by the equation $x = at/(2m+2+t^2)^{1/2}$ $\qquad$ (2)

then t has Student's distribution with $(2m+2)$ degrees of freedom. Use the Distribution to calculate the probability that $t_n \geq 2$, when $n = 2, 4$.

2*. Let $X_1, X_2, \ldots, X_n$ be independent observations from $N(\mu, \sigma^2)$ population and let $\bar{X}$ be the sample mean. If X' is one more observation independent of the previous ones, show that

$$\frac{(X - X')}{S_0}\left[\frac{n(n-1)}{n+1}\right]^{1/2} \sim t_{(n-1)}, \quad \text{or} \quad \left[\frac{(X'-X)}{\hat{S}}\left(\frac{n}{n+1}\right)^{1/2}\right] \sim t_{(n-1)}$$

where $\qquad S_0^2 = \Sigma(X_i - \bar{X})^2 = (n-1)\hat{S}^2.$

3*. If X is $\chi^2_{(n)}$, find $E(X^k)$. Deduce the moments expression for Student's $t_{(n)}$ and sample variance S^2. Hence evaluate Corr $(\bar{X}, T)$, where $\bar{X}$ is the mean of a sample of size n from $N(\mu, \sigma^2)$ and $T \sim t_{(n-1)}$.

4*. If $X \sim t_{(n)}$, show that $(n - \frac{1}{2})\ln[1 + (t^2/n)] \sim \chi^2_{(1)}$ (approximately).

You may assume that, for large n : $\dfrac{\Gamma(\frac{1}{2}n + \frac{1}{2})}{\Gamma(\frac{1}{2} + n)\sqrt{\frac{1}{2}n}} \sim \left(1 - \dfrac{1}{4n}\right).$

Exercise 22(a)

1. If X is $t_{(2)}$-distributed, find (a) $P\{X \geq 2\}$ (b) $P\{-\sqrt{2} \leq X \leq \sqrt{2}\}$.

2. Show that for A Student's t-distribution with n d.f., the m.g.f. does not exist but the absolute moment $\nu_r = E(|X|^r)$ exists. Find an expression for ν_r.

3. Let $\bar{X}$ and $\hat{S}^2$ be the usual sample mean and variance based on a sample of n independent observations from $N(\mu, \sigma^2)$ populations. Prove that

$$\text{Corr}\,(\bar{X}, \sqrt{n}\,(\bar{X} - \mu)/\hat{S}) = [\tfrac{1}{2}(n-3)]^{1/2}\,\Gamma[\tfrac{1}{2}(n-2)]/\Gamma[\tfrac{1}{2}(n-1)].$$

4. Let $I_x(a, b)$ represent the Incomplete Beta function defined by

$$I_x(a, b) = \frac{1}{B(a, b)}\int_0^x x^{a-1}(1-x)^{b-1}\,dx, \quad a > 0, b > 0.$$

Show that the distribution function $F(\cdot)$ of Student's t-distribution is given by

$$F(t) = 1, \tfrac{1}{2} I_x(\tfrac{1}{2}n, \tfrac{1}{2}), \text{ where } x = [1 + (t^2/n)]^{-1}.$$

Conclude that $[1 + (t^2/n)]^{-1}$ is $B_I(\frac{1}{2}n, \frac{1}{2})$.

5. If S is the S.D. based on n independent observations from a $N(\mu, \sigma^2)$-population, prove that, if $X \sim t_{(n)}$, then $V = C_n\ [2n/(X^2/2n]^{1/2}\ X$ has asymptotically the $N(0, 1)$ distribution where $C_n = E(S)/\sigma$.

[**Hint.** For n moderately large, S is approx. $N(\sigma C_n, \sigma^2/2n)$.

6. Let X, Y be independent $N(0, \sigma^2)$-variates. Show that $P\{[X+Y)^2/(X-Y)^2] < 4\} = 0.70$.

7. Let X_i, $1 \le i \le 5$ be i.i.d. $N(0, 1)$-variates. Determine a constant k such that the r.v. $k(X_1 + X_2)/(X_3^2 + X_4^2 + X_5^2)^{1/2}$ has t-distribution.

8. Let X_1 and X_2 be two independent $N(\mu, \sigma^2)$-variates. Obtain the distribution of

$Y = (X_1 + X_2 - 2\mu)/\sqrt{|X_1 - X_2|^2}$.

9. If X_i, $1 \le i \le n$ is a random sample of size n from $N(\mu, \sigma^2)$ distribution, show that

$U = (\overline{X} - \mu)\sqrt{n(n-1)}/\sqrt{\Sigma(X_i - \overline{X})^2}$ conforms to $t_{(n-1)}$-variate. If X is an additional observation from $N(\mu, \sigma^2)$, independent of X_i, show that

$$V = (X - \overline{X})\sqrt{n(n-1)}/S_0\sqrt{n+1} \sim t_{(n-1)}$$

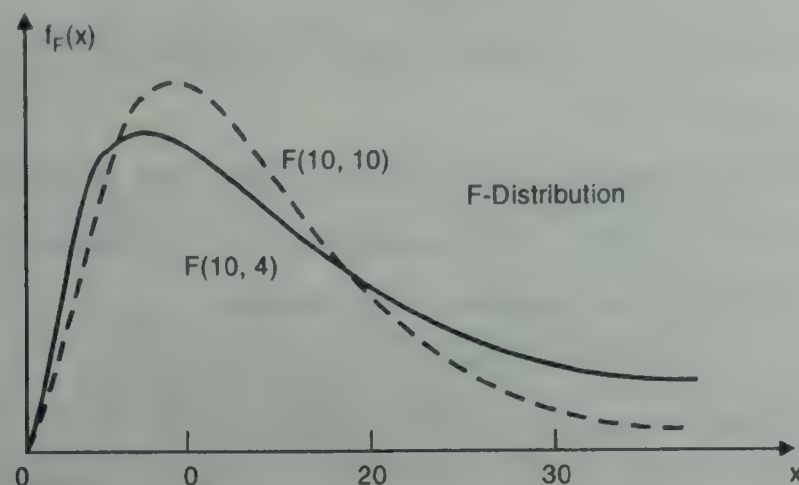
where S_0^2 is the sum of the squares of the deviations from $\overline{X}$.

SNEDECOR-FISHER F-DISTRIBUTION

22-40. *Definition*

A random variable X is said to possess F-distribution with m and n degrees of freedom, written $X \sim F\ (m, n)$, if its p.d.f is given by

$$f(x) = \frac{(m/n)^{\frac{1}{2}m}}{B(\frac{1}{2}m, \frac{1}{2}n)} \cdot \frac{(x)^{(m/2)-1}}{[1+(mx/n)]^{\frac{1}{2}(m+n)}}, \quad x > 0; \quad f(x) = 0,\ x \le 0. \qquad ...(1)$$



22-41. *Quotient of Two Independent Averaged Chi-squares*

If $Y \sim \chi^2_{(m)}$ and $Z \sim \chi^2_{(n)}$ are independent variates, then the variate

$$\boxed{U = \frac{Y/m}{Z/n} = \frac{n}{m}\frac{Y}{Z}} \qquad ...(1)$$

if $F\ (m, n)$-distributed.

Proof. Since Y and Z are indep. their joint elemental p.d.f. is

$$dP_1(y, z) = f_1(y)\, f_2(z)\, dy\, dz = \frac{e^{-y/2}\ y^{(m/2)-1}}{\Gamma(\frac{1}{2}m)\,(2)^{m/2}} \cdot \frac{e^{-\frac{1}{2}z} \,.\, z^{(n/2)-1}}{\Gamma(\frac{1}{2}n)\,.\,(2)^{n/2}}\, dy\, dz,\ 0 < y,\ z < \infty.$$

We now put $U = (nY/mZ)$, $V = Z$, and consider the transformations

$$u = ny/mz,\ v = z \ \Rightarrow\ y = (m/n)\, uv, \quad z = v$$

$$\therefore \qquad J = \frac{\partial(y, z)}{\partial(u, v)}\begin{vmatrix} \partial y/\partial u & \partial y/\partial v \\ \partial z/\partial u & \partial z/\partial v \end{vmatrix} = \begin{vmatrix} mv/n & \mu/n \\ 0 & 1 \end{vmatrix} = \frac{mv}{n}.$$

Thus, $dy\, dz = (mv/n)\, du\, dv$. Hence, the joint elemental p.d.f. of U & V is

$$dP_2\,(u, v) = \left(\frac{m}{n}\right)^{\frac{1}{2}m} \frac{e^{-\frac{1}{2}[1+(m/n)u]v}\ u^{(m/2)-1}\ v^{[(m+n)/2]-1}\, du\, dv}{(2)^{\frac{1}{2}(m+n)}\ \Gamma(\frac{1}{2}m)\,\Gamma(\frac{1}{2}n)},\ 0 < u,\ \ v < \infty.$$

To get the individual p.d.f. of U we integrate **out** unwanted v to get

$$dP(u) = \frac{(m/n)^{\frac{1}{2}m}\ u^{(m/2)-1} du}{(2)^{\frac{1}{2}(m+n)}\ \Gamma(\frac{1}{2}m)\,\Gamma(\frac{1}{2}n)} \int_0^{\infty} v^{[(m+n)/2]-1}\ e^{-\frac{1}{2}v[1+(mu/n)]} dv$$

$$= \left(\frac{m}{n}\right)^{\frac{1}{2}m} \frac{\Gamma[\frac{1}{2}(m+n)]}{(2)^{\frac{1}{2}(m+n)}\Gamma(\frac{1}{2}m)\Gamma(\frac{1}{2}n)} \frac{u^{(m/2)-1}\,du}{\left\{\frac{1}{2}[1+(mu/n)]\right\}^{(m+n)/2}} \qquad \text{[By § 0-11(a)]}$$

$$= \frac{(m/n)^{\frac{1}{2}m}\,u^{(m/2)-1}\,du}{B(\frac{1}{2}m, \frac{1}{2}n)\,[1+(mu/n)]^{(m+n)/2}},\ 0<u<\infty.$$

Thus, the p.d.f. of u, using dummy variable x, in place of dummy u, is given by

$$f(x) = \frac{(m/n)^{\frac{1}{2}m}\,x^{(m/2)-1}\,du}{B(\frac{1}{2}m, \frac{1}{2}n)\,[1+(mx/n)]^{(m+n)/2}},\quad 0<x<\infty.$$

Remark. § 22-41(1) may be treated as an alternate definition of $F(m, n)$.

***Cor.* 1.** If X is $F(m, n)$, then $(1/X)$ is $F(m, n)$.

Proof. Let $U \sim \chi^2_{(m)}$, $V \sim \chi^2_{(n)}$; then

$$X = \frac{U/m}{V/n},\ Y = \frac{1}{X} = \frac{V/n}{X/m}$$

The latter is the definition of $F(m, n)$ so $X^{-1} \sim F(n, m$

***Cor* 2.** If X is $t_{(n)}$, then X^2 is $\mathcal{F}(1, n)$.

Proof. Let $U \sim N(0, 1)$, $V \sim \chi^2_{(n)}$; then by definition

$$T = \frac{U}{\sqrt{(V/n)}} \Rightarrow T^2 = \frac{U^2}{V/n} = \frac{W/1}{V/n} = \mathcal{F}(1, n) \qquad [\text{because } W = U^2 \sim \chi^2_{(1)}.]$$

22-42. *Quotient of Two Independent Sample-scaled Variances*

If $X_1, X_2, ..., X_m$ is a random sample from $N(\mu_1, \sigma_1^2)$ and $Y_1, Y_2, ..., Y_n$ is another independent random sample from $N(\mu_2, \sigma_2^2)$, then

$$(\hat{S}_1^2/\sigma_1^2)/(\hat{S}_2^2/\sigma_2^2) \sim F(m-1, n-1).$$

Proof. Recall : $(\nu\hat{S}^2/\sigma^2) \sim \chi^2_{(\nu)}$ [§ 21-71]. Hence by § 22-41 (a)

$$F = \frac{(\chi_1^2/\text{rank})}{(\chi_2^2/\text{rank})} = \frac{(\nu_1\hat{S}_2^2/\sigma_1^2\,\nu_1)}{(\nu_2\hat{S}^2/\sigma_2^2\,\nu_2)} = \frac{(\hat{S}_2^2/\sigma_1^2)}{(\hat{S}_2^2/\sigma_2^2)}$$

is Snedecors $F(\nu_1, \nu_2)$ distributed.

Remark. § 22-42 is extremely useful for testing hypothesis $H_0: \sigma_1^2 = \sigma_2^2$.

***Example* 1.** Let X_1, X_2, be a random sample from $N(0, 1)$ and Y_1, Y_2 another random sample from a different distribution $N(1, 1)$. Find the distributions of the following :

(a) $(X_2 - X_1)/\sqrt{2}$ (b) $(X_2 + X_1)^2/(X_2 - X_1)^2$ (c) $(X_1 + X_2)/\sqrt{(X_1 + X_2)^2}$

(d) $\overline{X} + \overline{Y}$ (e) $(X_1 + X_2)/\sqrt{[(X_2 - X_1)^2 + (Y_2 - Y_1)^2]/2}$

(f) $[(Y_1 - Y_2)^2 + (X_1 - X_2)^2 + (X_1 + X_2)^2]/2$ (g) $(Y_1 + Y_2 - 2)^2/(Y_2 - Y_1)^2$.

Solution. Basic result. Linear combinations of indep. Normal variates is an appropriate Normal variate.

(a) Let $X = (X_2 - X_1)/\sqrt{2}, E(X) = 0, \text{Var}(X) = \frac{1}{2}(\text{Var } X_2 + \text{Var } X_1) = 1$, so $X \sim N(0, 1)$

(b) $U = (X_1 + X_2)/\sqrt{2} \sim N(0, 1), V = (X_2 - X_1)/\sqrt{2} \sim N(0, 1)$; $X_1 + X_2$ and $X_1 - X_2$ are independent

So
$$X' = \frac{(X_2 + X_1)^2}{(X_2 - X_1)^2} = \frac{U^2}{V^2} = \frac{\chi^2_{(1)}}{\chi^2_{(1)}} \sim F(1, 1).$$

(c)
$$X = \frac{X_1 + X_2}{|X_2 - X_1|} = \frac{(X_1 + X_2)/\sqrt{2}}{|(X_2 - X_1)/\sqrt{2}|} = \frac{U}{|V|}, \text{ say.}$$

$U \sim N(0, 1)$, and $V \sim N(0, 1)$ are independent. Hence $X \sim$ Chy (1, 0) : standard Cauchy

(d) $\bar{X} \sim N(\mu, \sigma^2/n)$, so in the present case $\bar{X} \sim N(0, 1/2), \bar{Y} \sim N(1, 1/2)$, hence by linear combination :

$$\bar{X} + \bar{Y} \sim N(1, 1).$$

(e) $(X_1 + X_2) \sim N(0, 2), (X_2 - X_1) \sim N(0, 2), (Y_2 - Y_1) \sim N(0, 2), (Y_1 + Y_2) \sim N(2, 2)$.

$$U = \frac{X_1 + X_2}{\sqrt{2}} \sim N(0, 1);\ V = \left(\frac{X_2 - X_1}{\sqrt{2}}\right)^2 + \left(\frac{Y_2 - Y_1}{\sqrt{2}}\right)^2 \sim \chi^2_{(2)}; \qquad [U, V \text{ are indep.}]$$

$$\therefore\ Z = \frac{X_1 + X_2}{\sqrt{[(X_2 - X_1)^2 + (Y_2 - Y_1)^2]/2}} = \frac{U}{\sqrt{V/2}} = \frac{N(0, 1)}{\sqrt{\chi^2_{(2)}/2}} \sim t_{(2)}.$$

(f) $W = \left(\frac{Y_1 - Y_2}{\sqrt{2}}\right)^2 + \left(\frac{X_1 - X_2}{\sqrt{2}}\right)^2 + \left(\frac{X_1 + X_2}{\sqrt{2}}\right)^2$ = sum of three $[N(0, 1)]^2$ variates ; $W \sim \chi^3_{(2)}$.

(g) $(Y_2 + Y_1 - 2)/\sqrt{2} \sim N(0, 1)$; hence

$$W = \frac{(Y_1 + Y_2 - 2)^2}{(Y_2 - Y_1)^2} = \frac{[(Y_1 + Y_2 - 2)/\sqrt{2}]^2}{[(Y_2 - Y_1)/\sqrt{2}]^2} = \text{ratio of two Indep. } \chi^2_{(1)}; \text{ hence } W \sim F(1, 1).$$

Example 2. Let $X_1, X_2, ..., X_n$ be a random sample from $N(\mu, \sigma^2)$. Define

$$\bar{X}_k = \frac{1}{k}\sum_1^k X_j, \qquad \bar{X}_{n-k} = \frac{1}{n-k}\sum_{k+1}^n X_j, \qquad \bar{X} = \frac{1}{n}\sum_1^n X_j$$

$$S_k^2 = \frac{1}{k-1}\sum_1^k (X_j - \bar{X}_k)^2,\quad S_{n-k}^2 = \frac{1}{n-k-1}\sum_{k+1}^\infty (X_i - \bar{X}_{n-k})^2,\quad S^2 = \frac{1}{n-1}\sum_1^n (X_i - \bar{X})^2.$$

Now find the distributions of the following :

(a) $\sigma^{-2}[(k-1)S_k^2 + (n-k-1)S_{n-k}^2]$ (b) $\frac{1}{2}(\bar{X}_k + \bar{X}_{n-k})$ (c) $(X_i - \mu)^2/\sigma^2$

(d) S_k^2/S_{n-k}^2 (e) $(\bar{X} - \mu)\sqrt{n}/S$.

Solution. (a) $(k-1)\,S_k^2/\sigma^2 \sim \chi^2_{(k-1)}$; $\quad (n-k-1)\,S^2_{n-k}/\sigma^2 \sim \chi^2_{n-k-1}$ [§ 21-70]

By Reproductive property of Indep. χ^2-variates, we conclude that their sum is a $\chi^2_{(n-2)}$ variate.

(b) $\overline{X}_k \sim N(\mu, \sigma^2/k)$; $\overline{X}_{n-k} \sim N[\mu, \sigma^2/(n-k)]$; Hence by Reproductive property

$$\overline{X}_k + \overline{X}_{n-k} \sim N\{2\mu, (\sigma^2/k) + \sigma^2/(n-k)\}.$$

Let $Z = \frac{1}{2}(\overline{X}_k + \overline{X}_{n-k})$; Then $\mathrm{E}(Z) = \mu$,

$$\text{Var}(Z) = \tfrac{1}{4}[(\sigma^2/k) + [\sigma^2/(n-k)]; \text{ whence } Z \sim N[\mu, n\sigma^2/4k(n-k)].$$

(c) Let $Z = [(X_j - \mu)/\sigma]^2 = [N(0,1)]^2$; so $Z \sim \chi^2_{(1)}$.

(d) $$Y = \frac{S_k^2}{S_{n-k}^2} = \frac{(k-1)\,S_k^2/\sigma^2(k-1)}{(n-k-1)\,S_{n-k}^2/(\sigma^2(n-k-1)} = \frac{Y_1/(k-1)}{Y_2/(n-k-1)} \sim F(k-1, n-k-1)$$

where $Y_1 \sim \chi^2_{(k-1)}$ and $Y_2 \sim \chi^2_{(n-k-1)}$.

(e) $$T = \frac{(\overline{X}-\mu)\sqrt{n}}{S} = \frac{(\overline{X}-\mu)}{\sigma/\sqrt{n}} = \sqrt{\frac{(n-1)\,S^2/\sigma^2}{(n-1)}} = \frac{U}{\sqrt{V/(n-1)}} \sim t_{n-1}$$

where $U \sim N(0, 1)$; $V \sim \chi^2_{(n-1)}$, so T is Student's $t_{(n-1)}$.

Exercise. Let $X_1, X_2, \ldots, X_n$ be a random sample from $N(0, 1)$ population. Define $\overline{X}_k$ and $\overline{X}_{n-k}$ as above (Example 2). Now assuming standard results, find the distribution of :

(a) $\frac{1}{2}(\overline{X}_k + \overline{X}_{n-k})$ (b) $k\overline{X}_k^2 + (n-k)\,\overline{X}_{n-k}^2$ (c) $\overline{X}_1^2/\overline{X}_2^2$ (d) X_1/X_2.

22-50. *Moments of F-distribution*

First Method. If X is $F(m, n)$, then its p.d.f. is

$$f(x) = Kx^{m/2-1}[1 + (mx/n)]^{-(m+n)/2},\ 0 \le x < \infty, \text{ where } K = (m/n)^{\frac{1}{2}m}\,[B(\tfrac{1}{2}m, \tfrac{1}{2}n)]^{-1}.$$

$$\mathrm{E}(X^r) = K\int_0^\infty \frac{x^{(m/2)+r-1}\,dx}{[1+(mx/n)]^{(m+n)/2}} = K\left(\frac{n}{m}\right)^{(m/2)+r}\int_0^\infty \frac{z^{(m/2)+r-1}\,dz}{(1+z)^{(m/n)/2}}, \qquad \left[\frac{m}{n}x = z\right]$$

$$= \left(\frac{n}{m}\right)^r \frac{B(\frac{1}{2}m + r, \frac{1}{2}n - r)}{B(\frac{1}{2}m, \frac{1}{2}n)} \qquad \left[\because \int_0^\infty \frac{x^{a-1}\,dx}{(1+x)^{a+b}} = B(a,b)\right]$$

$$\therefore \quad \mathrm{E}(X^r) = \left(\frac{n}{m}\right)^r \frac{\Gamma(\frac{1}{2}m+r)\,\Gamma(\frac{1}{2}n-r)}{\Gamma(\frac{1}{2}m)\,\Gamma(\frac{1}{2}n)} = \left(\frac{n}{m}\right)^r \frac{(\frac{1}{2}m)^{[r]}}{(\frac{1}{2}n-r)^{[r]}}. \qquad \ldots(1)$$

Second Method. If $U \sim \chi^2_{(m)}$ and $V \sim \chi^2_{(n)}$ are independent, then

$$X = (U/m) + (V/n) = (nU/mV) \sim F(m,n)$$

$$\therefore \quad \mathrm{E}(X^r) = \left(\frac{n}{m}\right)^r \mathrm{E}\left(\frac{U^r}{V^r}\right) = \left(\frac{n}{m}\right)^r \mathrm{E}(U^r)\,\mathrm{E}(V^{-r}) \qquad \ldots(2)$$

Now $\mathrm{E}(U^r) = 2^r\,\Gamma(\frac{1}{2}m+r)/\Gamma(\frac{1}{2}m); \mathrm{E}(V^{-r}) = 2^{-r}\,\Gamma(\frac{1}{2}n-r)/\Gamma(\frac{1}{2}n)$ ...[§21-20(1)]

Making substitution into (2), the result (1) follows.

22-51. *Moment Recurrence Formula*

Write $\mu'_r = \mathrm{E}(X^r)$; replace r by $r+1$ in §22-50 (1) to get

$$\mu'_{r+1} = \left(\frac{n}{m}\right)^{r+1} \frac{\Gamma(\frac{1}{2}m+r+1)\,\Gamma(\frac{1}{2}n-r-1)}{\Gamma(m/2)\,\Gamma(n/2)}.$$

Dividing this result by μ'_r we get

$$\mu'_{r+1} = \frac{n}{m}\frac{(\frac{1}{2}m+r)}{(\frac{1}{2}n-r-1)}\mu'_r = \frac{n(m+2r)}{m(n-2r-2)}\mu'_r. \qquad \ldots(1)$$

Mean and Variance. Putting $r = 0, 1$, the result (1) provides

$$\mu'_1 = \frac{n}{n-2},\quad \mu'_2 = \frac{n^2(m+2)}{m(n-2)(n-4)},\quad \mathrm{Var}(X) = \frac{2n^2(m+n-2)}{m(n-2)^2(n-4)},\quad n>4.$$

Coeff. of variation. $\dfrac{\sigma}{\mu} = \left[\dfrac{2(m+n-2)}{m(n-4)}\right]^{1/2}, \quad n>4.$

22-52. *Mode and Points of Inflexion*

If X is $F(m, n)$ its p.d.f. is given by

$$y = f(x) = Kx^{\frac{1}{2}m-1}(x+a)^{-(m+n)/2},\ [a = n/m \text{ and } K^{-1} = (m/n)^{n/2}\,B(\tfrac{1}{2}m, \tfrac{1}{2})],\ x>0 \qquad \ldots(1)$$

We take logarithms on both sides in (1) and then differentiate

$$\ln y = \ln K + (\tfrac{1}{2}m-1)\ln x - \tfrac{1}{2}(m+n)\ln(x+a).$$

$$y'/y = (\tfrac{1}{2}m-1)x^{-1} - \tfrac{1}{2}(m+n)(x+a)^{-1},\quad [y' = dy/dx] \qquad \ldots(2)$$

$$(y''/y) - (y'/y)^2 = \tfrac{1}{2}(m+n)(x+a)^{-2} - (\tfrac{1}{2}m-1)x^{-2}. \qquad \ldots(3)$$

Now $y' = 0 \Rightarrow \dfrac{m-2}{x} = \dfrac{m+n}{x+a} \Rightarrow x = \dfrac{a(m-2)}{(n+2)} = \dfrac{n}{m}\cdot\dfrac{m-2}{n+2} = M$ (say) ...(4)

Since $0 < x < \infty$, we must have $m > 2$. Putting for $x + a$ from (4) into (3) we get

$$\frac{y''}{y} = \frac{m+n}{2}\left[\frac{m-2}{(m+n)x}\right]^2 - \frac{m-2}{2x^2} = \frac{-(m-2)(2+m)}{2x^2(m+n)} < 0.$$

Thus, the value of x given in (4) gives mode of F (m, n)-distribution.

$$\therefore \quad \text{Modal value } M = \frac{n}{n+2}\cdot\frac{m-2}{m}(<1). \qquad [\text{each factor} < 1] \qquad \ldots(4)$$

Inflexion. Now put $y'' = 0$ in (3), and eliminate (y'/y) between (2) and (3) to get

$$\frac{1}{4}\left[\frac{m-2}{x} - \frac{m+n}{a+x}\right]^2 = \frac{1}{2}\left[\frac{m-2}{x} - \frac{m+n}{(a+x)^2}\right]$$

i.e. $[(n+2)x - a(m-2)]^2 = 2[a^2(m-2) + 2a(m-2)\,x - (n+2)x^2]$

or $(n+2)(n+4)x^2 - 2a(m-2)(n+4)\,x + a^2(m-2)(m-4) = 0.$ ...(5)

Observe that the sum and product of the two roots of this equation are

$x_1 + x_2 = 2M \Rightarrow x_2 - M = M - x_1 =$ distance of inflexion points from mode M.

$$x_1 x_2 = a^2(m-2)(m-4)/(n+2)(n+4).$$

Since $x_1, x_2 > 0$, $x_1\, x_1 > 0 \Rightarrow m > 4$. It follows that inflexion points exist only if $m > 4$. We now shift the origin to the mode M i.e. put

$$x = z + M = z + [a(m-2)/(n+2)] \qquad \text{...(6)}$$

Eq. (5) now reduces to

$$(n+2)(n+4)z^2 - 2a^2[(m+n)(m-2)/(n+2)] = 0$$

$$z = \pm a\sqrt{2}\,[(m+n)(m-n)]^{1/2}\,.[(n+4)(n+2)^2]^{1/2} \qquad \text{...(7)}$$

This also shows that the points of inflexion are equidistant from mode M.

22-60. Reciprocal Property

Let $X \sim F(m, n)$ and $Y \sim F(n, m)$, then for all a

(i) $P\{X \geq a\} = P(Y \leq (1/a)\}$ (ii) $P\{X \leq a\} + P\{Y \leq (1/a)\} = 1.$

Proof. (i) This result is trivially true, since $F(m, n) = 1/F(n, m)$

and $\{X \geq a\} = \{X^{-1} \leq a^{-1}\} = \{Y \leq 1/a\}.$

For another proof ; we record the p.d.f. of $F(m, n)$ as

$$dP_1(x) = Kx^{(m/2)-1}(n+mx)^{-(m+n)/2}dx,\ K = m^{m/2}\,.\,n^{n/2}\,.\,[B(\tfrac{1}{2}m, \tfrac{1}{2}n)]^{-1},\ x > 0 \qquad \text{...(1)}$$

Let $x = 1/y$, then $dx = -y^{-2}\,dy$ the range $(0, \infty)$, is mapped in the reverse $(0, \infty)$, hence absorbing the negative sign, the elemental p.d.f. of $Y = 1/X$ is

$$dP_2(y) = Ky^{-(1/2)m+1}(m+ny)^{-(m+n)/2}y^{(m+n)/2}y^{-2}dy = Ky^{(1/2)n-1}(m+ny)^{-(m+n)/2}dy \qquad \text{...(2)}$$

Result (2) is merely result (1) when m and n are interchanged.

Hence, if $X \sim F(m, n)$ then $Y \sim F(n, m)$. Consequently,

$$[X \geq a] = \{Y \leq a^{-1}\} \text{ which provides } P\{X \geq a\} = \{Y \leq a^{-1}].$$

(ii) By Negation-Rule, $P(X \geq a) + P(X \leq a) = 1$. Hence by (i), $P(X \leq a) + P(Y \leq a^{-1}) = 1$.

An equivalent expression for (ii) is $P(X \geq a) + P(Y \geq a^{-1}) = 1$.

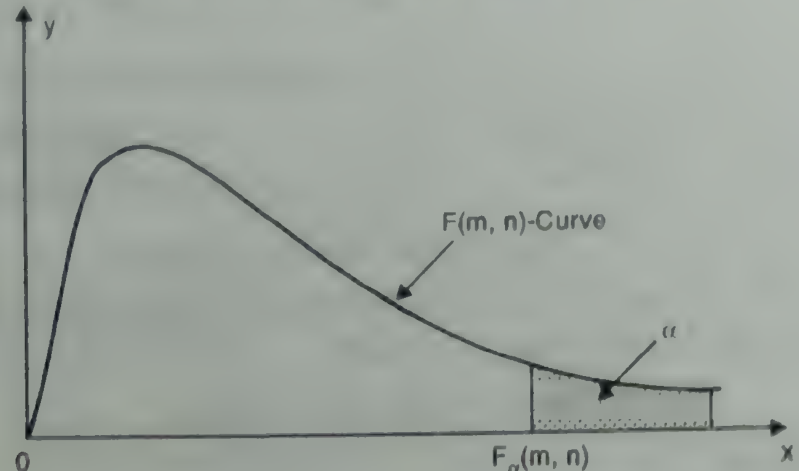
22-61. Reciprocal Relation Between the Upper and Lower α-points

If X is $F(m, n)$, we denote that point by $F_\alpha(m, n)$, in front of whose ordinate, lies a proportional α of the total area of distribution. Thus $P\{X \geq F_\alpha(m, n)\} = \alpha$.

Points $F_\alpha(m, n)$ and $F_{1-\alpha}(m, n)$ are called the upper α point and the lower α point respectively. The precise relation between them is

$$F_\alpha(n, m) \cdot F_{1-\alpha}(m, n) = 1. \quad \ldots(1)$$

Proof. If X is $F(m, n)$ and Y is $F(n, m)$; then by def. $Y = 1/X$.



$$P\{X > F_{1-\alpha}(m, n)\} = 1 - \alpha \quad \Rightarrow \quad P\{Y < 1/F_{1-\alpha}(m, n)\} = 1 - \alpha$$

$\therefore \quad P\{Y > 1/F_{1-\alpha}(m, n)\} = \alpha.$ (By Negation-Rule)

This means, area in front of $[F_{1-\alpha}(m, n)]^{-1}$ is α and Y is $F(n, m)$.

$$\therefore \quad F_\alpha(n, m) = 1/F_{1-\alpha}(m, n) \Rightarrow F_\alpha(n, m) \cdot F_{1-\alpha}(m, n) = 1.$$

Remarks. Let $m = n$.

If $\alpha = 0.50$, $F_\alpha(m, m) = Q_2$ hence $Q_2^2 = 1$.

If $\alpha = 0.25$, then $Q_1 \cdot Q_3 = 1$. If $\alpha = 0.10$, then $D_1 \cdot D_9 = 1$, and so on.

22-70. *Relation of F(m, n) to Beta Distributions*

(i) If $X \sim F(m, n)$, then $Z = [1 + (mX/n)]^{-1} \sim B_{I}(\frac{1}{2}m, \frac{1}{2}n)$

(ii) If $X \sim F(m, n)$, then $mX/n \sim B_{II}(\frac{1}{2}m, \frac{1}{2}n)$.

Proof. Since X is $F(m, n)$, its elemental p.d.f. is

$$dP_1(x) = Kx^{(1/2)m-1}[1 + (mx/n)]^{-(m+n)/2}\,dx, \; [K = (m/n)^{(1/2)m}/B(\tfrac{1}{2}m, \tfrac{1}{2}n)], \; x > 0. \quad \ldots(1)$$

Let $$z = \left(1 + \frac{mx}{n}\right)^{-1}, \text{ i.e. } x = \frac{m}{n}\left(\frac{1}{z} - 1\right), \; dx = -\frac{n}{m}\frac{1}{z^2}\,dz.$$

Substitutions into (1) provide

$$dP_2(z) = -K\left(\frac{n}{m}\frac{1-z}{z}\right)^{(m/2)-1} \cdot z^{(m+n)/2} \cdot \frac{n}{m} \cdot \frac{dz}{z^2} = \frac{-z^{(1/2)n-1}(1-z)^{(1/2)m-1}}{B(\frac{1}{2}m, \frac{1}{2}n)}\,dz. \quad \ldots(2)$$

We also observe that, when $x = 0$, $z = 1$; when $x \to \infty$, $z = 0$. Hence absorbing the negative sign, the p.d.f. of Z, as given by (2) is

$$f(z) = z^{(n/2)-1}(1-z)^{(m/2)-1}/B(\tfrac{1}{2}m, \tfrac{1}{2}n), \quad 0 < x < 1.$$

This shows that $Z \sim B_I(\frac{1}{2}m, \frac{1}{2}n)$.

(ii) By definition, $F = \dfrac{\chi_1^2 / m}{\chi_2^2 / n} \Rightarrow \dfrac{mF}{n} = \dfrac{\chi_1^2}{\chi_2^2} = \dfrac{\text{gam}\,(\frac{1}{2}, m)}{\text{gam}\,(\frac{1}{2}, n)} = B_{II}\left(\dfrac{m}{2}, \dfrac{n}{2}\right)$ [§19-33(a)]

22-71. *Limiting Relation of F(m, n) to Chi-square Distribution*

As $n \to \infty$, $m\, F$ tends to be distributed as $\chi^2_{(m)}$.

Proof. The elemental p.d.f. for F variate is

$$dP_1(F) = \frac{m^{m/2} \,.\, n^{n/2}}{B(\frac{1}{2}m, \frac{1}{2}n)} \cdot \frac{F^{(m/2)-1} dF}{(n + mF)^{(m+n)/2}}, \quad 0 < F < \infty.$$

Setting $mF = x$, $B(\frac{1}{2}m, \frac{1}{2}n) = \Gamma(\frac{1}{2}m)\,\Gamma(\frac{1}{2}n) / \Gamma(\frac{1}{2}m + \frac{1}{2}n)$; this becomes

$$dP_2(x) = \frac{x^{(1/2)m-1}\, dx}{\Gamma(\frac{1}{2}m)\,(2)^{m/2}} \left(1 + \frac{x}{m}\right)^{-(m+n)/2} \frac{\Gamma(\frac{1}{2}n + \frac{1}{2}m)}{\Gamma(\frac{1}{2}n)\,(\frac{1}{2}n)^{m/2}} \qquad \text{...(1)}$$

We now recall the standard limits

$$\lim_{n \to \infty} \left(1 + \frac{a}{n}\right)^{bn + c} = e^{ab} \text{ (Euler's limit)}; \quad \lim_{N \to \infty} \frac{\Gamma(N + k)}{N^k \Gamma(N)} = 1 \qquad \text{...(2)}$$

In (2), we put $a = x$, $b = -\frac{1}{2}$, $c = -\frac{1}{2}m$; $N = \frac{1}{2}n$, $k = \frac{1}{2}m$; taking limits in (1) as $n \to \infty$ and utilize (2) with new values to get

$$dP(x) = x^{(m/2)-1}\, e^{-x/2}\, dx / \Gamma(\tfrac{1}{2}m)\,(2)^{m/2}, \; 0 < x < \infty. \qquad \text{...(3)}$$

This is obviously the p.d.f. for a $\chi^2_{(m)}$-variate.

Cor. If $n \to \infty$, and $m = 1$ then $\sqrt{F} \sim N(0, 1)$.

Since $mF \to \chi^2_{(m)}$, it follows that $\sqrt{F} \to \sqrt{\chi^2_{(1)}} = N(0, 1)$.

22-80. *Fisher's Z-distribution*

Fisher's Z is defined by $Z = \frac{1}{2} \ln_e F$, where F is Snedecor's F-statistic. The distribution of Z is given by the usual transformation law :

$h(z) = p(F)\; | \partial F / \partial z | \qquad (F = e^{2z})$

$$= \frac{(m/n)^{(m/2)}\, (e^{2z})^{(m/2)-1}\, 2e^{2z}}{B(\frac{1}{2}m \frac{1}{2}n)\, [1 + (me^{2z}/n)]^{(m+n)/2}} = \frac{2(m/n)^{(m/2)}}{B(\frac{1}{2}m, \frac{1}{2}n)} \cdot \frac{e^{mz}}{[1 + (me^{2z}/n)]^{(m+n)/2}}, \quad -\infty < z < \infty.$$

$h(z)$ is the p.d.f. of Fisher's Z with m and n degrees of freedom.

Remark. The relation $Z = \frac{1}{2} \ln_e F$, implies that Z and F will bear the same Statistical applications.

22-81. Moment Generating Function of Z-distribution

$$M(t:Z) = \mathrm{E}(e^{tZ}) = \mathrm{E}(e^{t/2 \ln F}) = \mathrm{E}[e^{\ln(F)^{t/2}}] = \mathrm{E}(F^{t/2}) = \mu'_{t/2}.$$

Here $\mu'_{t/2}$ is the $(t/2)$th order simple moment of Snedecor F-statistic.

$$M(t:Z) = \mathrm{E}(F^{t/2}) = \left(\frac{n}{m}\right)^{t/2} \frac{\Gamma[\frac{1}{2}(m+t)].\Gamma(\frac{1}{2}(n-t)]}{\Gamma(\frac{1}{2}m).\Gamma(\frac{1}{2}n)}. \qquad ...(1)$$

22-82. Worked-out Problems

***Example* 1.** Show that F-distribution is highly positively skewed. Prove that if X is $F(n, n)$ its median is at $X = 1$ and its quartiles Q_1 and Q_3 satisfy the condition $Q_1 Q_3 = 1$.

Solution. Karl Pearson's coefficient of Skewness is given by

$$S_k = [(\text{Mean} - \text{Mode})/\text{S.D.}] > 0$$

because Mean $= [n/(n-2)] > 1$ and Mode $= [(n-2)/(n+2)] < 1$. Hence F-distribution is highly positively skewed.

When $m = n$, the elemental probability law of X is

$$f(x) = Kx^{(n/2)-1} / (1+x)^n \quad [K^{-1} = B(\tfrac{1}{2}n, \tfrac{1}{2}n)],\ 0 < x < \infty.$$

If Q_r is the rth quartile, then its definition provides

$$\int_0^{Q_r} f(x)\,dx = \int_0^{Q_r} \frac{Kx^{(n/2)-1}}{(1+x)^n}\,dx = \frac{r}{4} \qquad ...(1)$$

Let $x = 1/z$, then $dx = -z^{-2}\,dz$, $x = 0 \Rightarrow z \to \infty$; $x = Q_r \Rightarrow z = 1/Q_r = q_r$, say. Making substitutions into (1), we get

$$\int_{q_r}^{\infty} \frac{K z^{(n/2)-1}}{(1+z)_n}\,dz = \frac{r}{4} \text{ or } \int_0^{q_r} f(x)\,dx = 1 - \frac{r}{4}. \qquad ...(2)$$

Using $r = 2$, Eqs. (1) and (2) give

$$\int_0^{q_2} f(x)\,dx = \frac{1}{2} = \int_0^{Q_2} f(x)\,dx.$$

This gives $q_2 = Q_2$ or $Q_2^2 = 1 \Rightarrow Q_2 = +1$. Now put $r = 1$ into (1) and $r = 3$ into (2) to get

$$\int_0^{q_3} f(x)\,dx = \tfrac{1}{4} = \int_0^{Q_1} f(x)\,dx.$$

This gives $q_3 = Q_1 \Rightarrow Q_1 Q_3 = 1$.

Note. The solution follows trivially from Reciprocal Relations §22-61.

***Example* 2.** Let $X_i \sim N(i, i^2)$, $i = 1, 2, 3$ be independent variates. Using only X_1, X_2, X_3 give an example of a statistic that has the distribution (a) $\chi^2_{(3)}$, (b) $F(1, 2)$, (c) $t_{(2)}$.

Solution. Observe $Z_i = [(X_i - i)/i] \sim N(0, 1)$. Now

(a) $Z_1^2 + Z_2^2 + Z_3^2 = (X_1 - 1)^2 + [(X_2 - 2)/2]^2 + [(X_3 - 3)/3]^2 = Y \sim \chi^2_{(3)}$.

(b) $Z_1^2 \sim \chi^2_{(1)}$; $(Z_2^2 + Z_3^2) \sim \chi^2_{(2)}$; so $F(1, 2) = \dfrac{Z_1^2/1}{(Z_2^2 + Z_3^2)/2} = \dfrac{2(X_1 - 1)^2}{[\frac{1}{2}(X_2 - 2)]^2 + [\frac{1}{3}(X_3 - 3)]^2}$.

(c) $t_{(2)} = \dfrac{N(0, 1)}{\sqrt{\chi^2_{(2)}/2}} = \dfrac{Z_1}{[(Z_2^2 + Z_3^2)/2]^{1/2}} = \dfrac{\sqrt{2}(X_1 - 1)}{\left\{[\frac{1}{2}(X_2 - 2)]^2 + [\frac{1}{3}(X_3 - 3)]^2\right\}^{1/2}}$

Problems with Solutions Provided at the End of the Text

1*. If the variate X is $B_1(a, b)$ then the variate Z :

$$Z = \frac{X/a}{(1-X)/b} = \frac{bX}{a(1-X)} = \frac{X/\mu}{(1-X)/(1-\mu)}\left(\mu = \frac{a}{a+b}\right)$$ is distributed as $F(2a, 2b)$.

2*. (i) If $X \sim t_{(n-2)}$, then $Y \sim F(n-2, n-2)$ when $X = \frac{1}{2}(n-2)^{1/2}(1-Y)Y^{-1/2}$.

(ii) If X is $F(n, n)$, then $Y \sim t_{(n)}$, when $Y = \frac{1}{2}\sqrt{n}\,[X^{1/2} - X^{-1/2}]$.

3*. Given that : $P\{F(10, 12) > 2.753\} = 0.05$, $P\{(1, 12) > 4.747\} = 0.05$.

Find $p_1 = P\{F(12, 10) > 1/2.753\}$, $p_2 = P\{-\sqrt{4.744} < t_{(12)} < \sqrt{4.747}\}$.

4*. Let $X \sim F(2, n)$, $(n \geq 2)$. Show that $P\{X \geq k\} = [1 + (2k/n)]^{-n/2}$.
Deduce the significance level of F corresponding to the significance level of probability p.

Exercise 22(b)

1. If $X_1, X_2, \ldots, X_m, X_{m+1}, \ldots, X_{m+n}$ are i.i.d. $N(0, \sigma^2)$ variates, find the p.d.f. of $\sum_{i=1}^{m} X_i^2 \Big/ \sum_{i=m+1}^{m+n} X_i^2$.
2. Show how the probability points of $F(m, n)$ can be obtained from those of $F(n, m)$.
3. If X and Y are independent Expo (λ)-variates, show that $Z = (X/Y)$ has $F(2, 2)$ distribution. Hence or otherwise show that Z and Z^{-1} are identically distributed.
4. If $X \sim \chi^2_m$ and $Y \sim \chi^2_n$ are independent variates, show that $U = X + Y$, $V = nX/mY$ are independently distributed; and identify their distributions.
5. If $X \sim$ bin (n, p) and $Y \sim F\,[2k, 2(n-k+1)]$ show that,
$$P\{X < k\} = P\{Y > p(n-k+1)/kq\}.$$
6. If X_i, $1 \leq i \leq n$ are i.i.d. $N(0, \sigma^2)$ variates, find the p.d.f. of $W = X_1/[\Sigma X_i^2/n]^{1/2}$.
Why W does not follow t-distribution ?
7. Deduce the moments of $F(m, n)$ from those of $\chi^2_{(k)}$.
8. If $X \sim F(1, n)$, show that $(n - \frac{1}{2})\,\ell n[1 + (X/n)] \sim \chi^2_{(1)}$, for large n.
9. If $X \sim F(m, n)$, show that $W = (mX/n)/[1 + (mX/n)] \sim B_1(\frac{1}{2}m, \frac{1}{2}n)$. Deduce the Var (X) from p.d.f. of W.

10. Let $X_j \sim \chi^2_{r_j}, 1 \le j \le 3$ be independent variates. Prove that

$$\frac{X_1 / R_1}{X_2 / r_2} \text{ and } \frac{X_3 / r_3}{(X_1 + X_2)/(r_1 + r_2)} \text{ are independent } F\text{-variables.}$$

11. A random sample of n observations $(x_i, y_i)\ i = 1, 2, \ldots, n$ is drawn from a BVN $(\mu_1, \mu_2, \sigma^2, \sigma^2, \rho)$

Show that $W = [(1-\rho)S_u^2 / (1 + \rho S_v^2] \sim F(n-1, n-1)$

where $u_i = x_i + y_i,\ v_i = x_i - y_i,\ n\bar{u} = \Sigma u_i,\ n\bar{v} = \Sigma v_i$

$$(n-1)\, S_u^2 = \Sigma(u_i - \bar{u})^2, (n-1) S_v^2 = \Sigma(v_i - \bar{v})^2.$$

Statistics will prove anything; even the truth (or even the existence of God). [N. Moynihan]

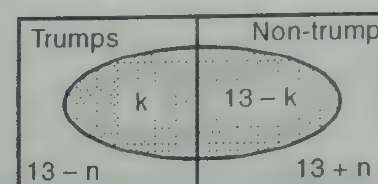
Each betrayal begins with trust. (Phish)

Solutions to Starred Problems

Chapter 1 : Elementary Probability : Basic Concepts

Sec. 1-41. Page 10

1*. East and West together have 26 cards, and these cards include $(13 - n)$ cards of the suit S. East (say) has k cards out of these $(13 - n)$ cards of suit S and $(13 - k)$ out of the remaining $(13 + n)$ cards. By sequential counting, the required probability p is



$$p = 2 \times \binom{13-n}{k}\binom{13+n}{13-k} \Big/ \binom{26}{13}.$$

The factor 2 signifies the fact that either of the East or West has exactly k trumps.

2*. Here $S = \{1, 2, ..., 17\}$. The set A of numbers which are divisible by 3 or 7 is given by $A = \{3, 6, 9, 12, 15, 7, 14\}$. Hence by Laplace definition $P(A) = 7/17$.

3*. (i) In a leap year, there are 52 weeks + 2 days. These two *consecutive* days have the following possibilities :

Ω = {(Sun, Mon), (Mon, Tue) (Tue, Wed), (Wed, Thur), (Thur, Fri), (Fri, Sat), (Sat, Sun)}.

Of these seven possibilities, only two combinations yield Sundays. Thus to get a Sunday in these situations, the probability is 2/7.

(ii) In a non-leap year, there are 52 weeks + 1 day. This day can be any one of the seven-days of the week. Thus, for this day to be Sunday, the probability is 1/7.

4*. Five tickets can be chosen out of 50 in $\binom{50}{5}$ ways; so $n(S) = \binom{50}{5}$. Since $x_3 = 30$, tickets x_1 and x_2 must come out of the set $\{1, 2, ..., 29\}$ and tickets x_4 and x_5 must come out of the set $\{31, 32, ..., 50\}$. By the sequential counting method, tickets x_1, x_2, x_4, x_5 can be choosen in $\binom{29}{2}\binom{20}{2}$ ways. By Laplace definition, $p = \binom{29}{2}\binom{20}{2}\Big/\binom{50}{5}$.

5*. Since a Bridge-hand consists of 13 cards, the total number of exhaustive cases is $N = \binom{52}{13}$. Consider a particular suit, say Diamonds (D). Nine cards of D can be selected out of 13 in $\binom{13}{9}$ ways and the rest of 4 cards can be selected out of 39 non-diamonds in $\binom{39}{4}$ ways. By Sequential-counting law, 13 cards containing 9 diamonds can be had in $f = \binom{13}{9}\binom{39}{4}$ ways. Hence

$$P\,\{4 \text{ non-diamonds and 9 diamonds}\} = \binom{13}{9}\binom{39}{4}\Big/\binom{52}{13}. \qquad ...(i)$$

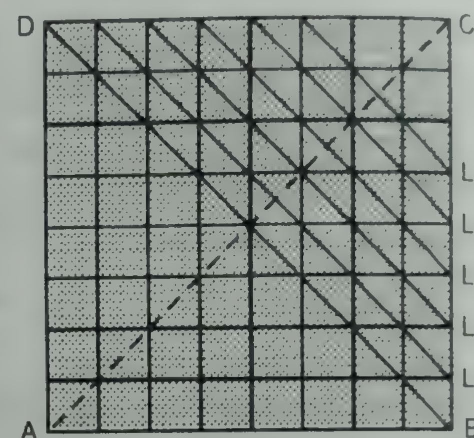
In a pack, there are four suits : Diamonds, Hearts, Clubs, Spades, each meeting the situation in (i). Hence, the required probability is

$$p = 4\binom{13}{9}\binom{39}{4}\Big/\binom{52}{13}.$$

6*. $p_1 = \dfrac{12}{12}\cdot\dfrac{8}{12}\cdot\dfrac{4}{12} = \dfrac{2}{9}$; $\quad p_2 = \dfrac{12}{12}\cdot\dfrac{8}{11}\cdot\dfrac{4}{10} = \dfrac{16}{55}$. [Supply reasoning]

Other methods are also possible.

7*. In a chess-board, there are $8 \times 8 = 64$ small squares. And any three small squares can be chosen at random in $N = \binom{64}{3} = 64 \times 31 \times 21 = 41664$ ways.



To choose 3 small squares so that they lie on a diagonal line we observe that they lie on lines such as L_1, L_2, ..., L_5, BD and lines on the other sides of BD, and an equal number of opportunities on the either side of the other main diagonal AC. Hence, by the *rule of sum*, the total number f of favourable cases is

$$f = 2\binom{8}{3} + \left\{\binom{7}{3} + \binom{6}{3} + \binom{5}{3} + \binom{4}{3} + \binom{3}{3}\right\}$$
$$= 2 \times 46 + 4(35 + 20 + 10 + 4 + 1) = 112 + 280 = 392.$$

By Laplace Definition : $p = f/N = 392/41664 = 7/744$.

8*. Six strings aA, bB, cC, dD, eE, fF are shown in the problem. Now a can be tied to any one of the five upper ends b, ..., f. One join is ab. Then c can be tied to any one of the three upper ends (d, e, f), one join is cd and finally e can be tied only to one point f. Thus three are $5 \times 3 \times 1 = 15$ different ways of joining in pairs the upper ends of the strings. Likewise, there are 15 different ways of joining the lower ends A, B, ..., F. By rule of sequential counting, the total number of ways in which ties are carried out is $N = 15 \times 15 = 225$; all tying together being equally likely.

Favourable Cases : Suppose the upper ends joined yield ab, cd, ef. In order to obtain a single ring, it is necessary that A be tied to C or D or E or F (this gives 4 possibilities for A). Now if A is joined to C, then B will have to be joined to ends E or F (this gives 2 possibilities). This leaves only two loose ends, say, D and F, which must be joined together. Counting all possibilities, we see that for each of the 15 ways of joining A, B, ..., F there are $4 \times 2 \times 1 = 8$ ways of joining a, b, ..., f so that the favourable outcomes are to $15 \times 8 = 120$. Hence

$$P\{\text{a single ring is formed}\} = t/N = 120/225 = 8/15$$

$$P\{\text{at least one ring is formed}\} = 1 - P(\phi) = 1.$$

Comments. If there are $2n$ strings, then

$$N = [(2n-1)(2n-3) \ldots 5.3.1]^2 \; ; \; t = [(2n-1) \ldots 3.1]\,[(2n-2)(2n-4) \ldots 4.2]$$

$$\therefore \quad p = \frac{(2n-2)(2n-4)\ldots 4.2}{(2n-1)(2n-3)\ldots 5.3} \qquad [n = 3 \text{ yields the above case}]$$

9*. Let M_r denote that r people have same birth month. Trivially, we may omit event M_1. The event space is thus expressible : $E = \{M_4, M_3, M_2, M_1, M_0\}$.

Let J_1 = {first person has Jan. as birth month}, F_1 = {first person has Feb. as birth month}, etc.

$$\therefore \quad P(J_1, J_2, J_3, J_4) = P(J_1) \ldots P(J_4) = (1/12)^4.$$

So $\quad P\{M_4\} = 12 \times (1/12)^4 = (1/12)^3$ [$\because$ There are 12 different months]

Again, $P\{J_1, J_2, J_3, F_4\} = (1/12)^4$, [4th born in Feb. etc.]

Two months out of 12 can be chosen in $\binom{12}{2}$ ways.

And arrangements of 4 objects of two types is $4!/(3!\ 1!)$.

From $(J_1, J_2, J_3, F_4) = \{F_1, F_2, F_3, J_4\}$, we conclude that

$$P\{M_3\} = 2\binom{12}{2}\frac{4!}{3!\,1!}\left(\frac{1}{12}\right)^4 = \frac{44}{(12)^3}.$$

From $P(J, J, F, M\} = (1/12)^4$; [$JJFM$); ($JFFM$), ($JFMM$), 3 possibilities] and proceeding as above

$$P\{M_2\} = 3\binom{12}{3}\cdot\frac{4!}{2!\ 1!\,1!}\cdot\left(\frac{1}{12}\right)^4 = \frac{660}{(12)^3}.$$

$$P\{M_1\} = \binom{12}{2}\cdot\frac{4!}{2!\ 2!}\cdot\left(\frac{1}{12}\right)^4 = \frac{33}{(12)^3}. \qquad [(JJFF), (FFJJ), \ldots.]$$

$$P\{M_0\} = \binom{12}{2}\cdot\frac{4!}{1!\ 1!\,1!\,1!}\cdot\left(\frac{1}{12}\right)^4 = \frac{990}{(12)^3} = \frac{990}{1728}.$$

We may note that : $\displaystyle\sum_{k=0}^{4} P(M_k) = \frac{(1+44+660+33+990)}{(12)^3} = 1.$

(a) $P(E_2) = P\{M_2\} = 660/(12)^3 = 0.38$.

(b) $P\{L_2\} = (1 + 44 + 660 + 33)/(12)^3 = 0.43$.

(c) $P\{D\} = 990/1728 = 0.57$.

10*. (a) If $n > N$, $p = 0$ trivially; hence we assume that $n \le N$. The first ball can be placed in any one of N boxes, the second into any one of the remaining $(N-1)$ boxes and so on. Thus, in all there are $N^{(n)} = N(N-1)(N-2) \ldots (N-n+1)$ different ways to place n balls into N boxes, no box receiving more than one ball.

Since each of the n balls can go into anyone of the N boxes, in all, there are N^n different ways of distributing the balls into the boxes. Thus, the required probability, by Laplace definition, is $p = N^{(n)}/(N)^n$.

(b) Let the jth ball be in ith box. Then there are $n - 1$ more balls to distribute in the N boxes with no restrictions on where they go. Obviously, there are $(N)^{n-1}$ different ways of distributing these $n - 1$ balls into the N boxes. Hence $p = (N)^{n-1}/(N)^n = 1/N$.

11*. Let the consecutive numbers be $x, x + 1, x + 2, \ldots, x + 3n - 1$. We arrange them in three rows, each consisting of n numbers as follows :

R_1 : $x, x + 3, x + 6, \ldots, x + 3n - 3$.

R_2 : $x + 1, x + 4, x + 7, \ldots, x + 3n - 2$.

R_3 : $x + 2, x + 5, x + 8, \ldots, x + 3n - 1$.

12*.
$$P\{\text{all are black}\} = \binom{8}{3}\Big/\binom{20}{3} = \frac{14}{285};\quad P(2b, w) = \binom{8}{2}\binom{9}{1}\Big/\binom{20}{3} = \frac{21}{95}$$

$$P(b, r, w) = \binom{8}{1}\binom{3}{1}\binom{9}{1}\Big/\binom{20}{3} = \frac{18}{95};\quad P(b, r, w) = \frac{1}{3!}\,P(BRW) = \frac{3}{95}$$

Independently, $$P(BRW) = P(B)\,P(R\,|\,B)\,P(W\,|\,BR) = \frac{8}{20}\cdot\frac{3}{19}\cdot\frac{9}{18} = \frac{3}{95}.$$

13*. Let A be fixed at any point on the ring, then B can occupy any of the 11 remaining places. Of these 11 places, there are only 2 places (B_1 and B_2 say) such that the gap between A and B is of three places. Hence $p = 2/11$. [Sample space is assumed to be equiprobable].

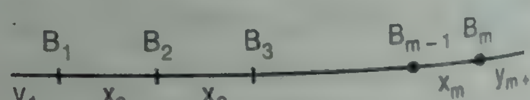
14*. For circular arrangements, we fix one person for counting to commence. This gives $N = (n + 1)\,!$. Now two groups consisting of k persons and $(n - k)$ persons between A and B can be permuted in $k!$ and $(n - k)!$ and k persons can be selected in ${}^nC_k$ ways. A and B can also interchange their positions in 2! ways. Hence, the favourable number of cases leading to the stated geography $f = k!\,(n - k)!\,{}^nC_k \,.\, 2! = 2(n!)$. Thus $p = f/N = 2(n!)/(n + 1)! = 2/(n + 1)$.

15*. The entire composition of the event space Ω is composed of 24 outcomes :

{*abcd, abdc, acbd, acdb, adbc, adcb, bacd, badc, bcad, dcda, bdac, bdca, cabd, cadb, cbad, cbda, cdab, cdba, dabc, dacb, dbac, dbca, dcab, dcba*}.

The favourable outcomes include the 6 in the first block, none in the 2nd and 3rd blocks, and the first two in the 4th block; totalling in all 8 units. Thus $p = 8/24 = 1/3$.

16*. *Gap Method.* Let y_1 be the number of red balls ere B_1 (first black ball) is drawn, x_2 be the number of red balls drawn ere B_2 is drawn and so on and suppose y_{m+1} red balls remain after B_m is drawn. Then, n being total red balls, we have



$$y_1 + x_2 + x_3 + \ldots + x_m + y_{m+1} = n, \quad y_1 \geq 0,\ y_{m+1} \geq 0,\ x_j > 0.$$

Put $1 + y_1 = x_1$, $y_{m+1} + 1 = x_{m+1}$ then this equation becomes,

$$x_1 + x_2 + \ldots + x_m + x_{m+1} = n + 2, \quad x_j > 0,\ j = 1, 2, \ldots, m + 1.$$

The number of solutions, in integers, of this equation is

$$f = \binom{n+2-1}{n+1-1} = \binom{n+1}{m}$$

As $N = \binom{m+n}{m}$, so $p = \frac{f}{N} = \binom{n+1}{m} \Big/ \binom{m+n}{m}$.

17*. Suppose that the extraction continues until there is left in the urn only one ball. The colour of this last ball is obviously the same as the single colour remaining in the urn. Thus, the probability that the single colour remaining is white equals probability that the last ball in the urn is white.

The number f of favourable cases for the last ball to be white is obtained by considering that among the $w + b - 1$ balls, $w - 1$ balls were white; hence

$$f = \binom{w+b-1}{w-1}, \text{ and so } p = \binom{w+b-1}{w-1} \Big/ \binom{w+b}{w} = \frac{w}{w+b}.$$

Comments. The result is obvious if we attend to the fact that the last ball to be extracted is every bit as random as the first ball and probability that the first ball drawn out is white, is $w/(w + b)$.

18*. The even number of cards drawn means that either 2, or 4, ..., or 52 cards are drawn. Hence by Rule of Sum

$$n(S) = \binom{52}{2} + \binom{52}{4} + \ldots + \binom{52}{52} = 2^{51} - 1 \qquad \ldots(1)$$

This sum is obtained by putting $x = 1$, $n = 52$ in the identity

$$(1 + x)^n + (1 - x)^n = \sum_{r=0}^{n} \binom{n}{r} [1 + (-1)^r]\, x^r \quad \Rightarrow \quad 2^{52} = 2 + 2\left[\binom{52}{2} + \binom{52}{4} + \ldots + \binom{52}{52}\right]$$

Let $n(R_j, B_j)$ denote the number of j red and j black cards drawn; then, by Sequential Counting, $n(R_j, B_j) = \binom{26}{j}\binom{26}{j}$ and so

$$p = \sum_{j=1}^{26} \binom{26}{j}\binom{26}{j} \Big/ n(S) \qquad \ldots(2)$$

Put $n = 26$ in the Hypergeometric Identity,

$$\sum_{r=0}^{n} \binom{n}{r}\binom{n}{n-r} = \sum_{r=0}^{n} \binom{n}{r}\binom{n}{r} = \binom{2n}{n}; \text{ to get } 1 + \sum_{r=1}^{26} \binom{26}{r}^2 = \binom{52}{26} \qquad \ldots(3)$$

From (1), (2) and (3) we obtain,

$$p = \left[\binom{52}{26} - 1\right] \Big/ (2^{51} - 1).$$

Note that $N = N_1 \cdot N_2 \cdot N_3 \cdot N_4$ and N_1 are not independent, for $P(N \mid N_1) \neq P(N)$.

19*. Let x_1 be the number of persons on the left of A and x_2 be the number of persons to the right of B. Then

$$x_1 + k + x_2 = n, \text{ i.e. } x_1 + x_2 = n - k, \; x_1 \geq 0, x_2 \geq 0.$$

The number of solutions, in integers, of this equation is

$$N_1 = \binom{n-k+2-1}{n-k} = n + 1 - k.$$

As A and B can interchange their positions in 2! ways, the favourable cases are $f = 2N_1 = 2(n+1-k)$. Further, A can occupy any of $(n+2)$ slots and then B can occupy any of the remaining $(n+1)$ slots, the total number of available slots is $N = (n+2)(n+1)$. Hence

$$p = f/N = 2(n+1-k)/(n+2).(n+1).$$

Sec. 1-56. Page 32

1*. (a) Suppose $A \cap B = \varnothing$ (disjoint events), then $A \subseteq \bar{B}$ (draw Figure), and so $P(A) \le P(\bar{B})$.

This is obviously wrong because $P(A) > P(B)$.

(b) Since $AB \subseteq A$, so $P(AB) \le P(A) = 0$. It follows that $P(AB) = 0$, which is true.

(c) Suppose $\Omega = \{e_1, e_2, ..., e_6\}$, $p_i = P(e_i)$ and assign the probabilities

$$p_1 = \theta_1,\ p_2 = 5\theta,\ p_3 = 2\theta,\ p_4 = p_5 = p_6 = 4\theta,\ (\Sigma p_i = 1 \Rightarrow \theta = 1/20 = 0.05).$$

Define the events $A = \{e_1, e_3, e_4\}$, $B = \{e_1, e_3, e_6\}$, then $AB = \{e_1, e_3\}$. Now,

$$P(A) = 0.35,\ P(B) = 0.35,\ P(AB) = 0.15,\ p^2 = (0.35)^2 = 49/400,\ P(AB) = 15/100;$$

$$P(AB) < p^2 = 60 \le 49 \text{ which is false. Thus } P(AB) \le p^2, \text{ is untrue.}$$

Note. $P(A) = P(B)$, but $A \neq B$.

(*d*) Let Ω and p_j be as in part (*c*). We define the events : $A = \{e_4\}$, $B = \{e_1, e_2, e_3, e_4, e_5\}$.

Now $\bar{B} = \{e_6\}$ and $\bar{A} = \{e_1, e_2, e_3, e_4, e_5, e_6\}$. Obviously $P(A) = 0.20 = P(\bar{B})$, but $\bar{A} \neq B$.

2*. Let D_r denote rth-numbered draw. Now event A occurs if odd-numbered balls are chosen on D_1 and D_3 (event A_1) or on D_3 and D_5 (event A_1) or on D_3 and D_5 (event A_3). As $A = A_1 \uplus A_2 \uplus A_3$, we have

$$P(A) = P(A_1 \uplus A_2 \uplus A_3) = P(A_1) + P(A_2) + P(A_3) = 3P(A_1). \quad \text{[by symmetry]}.$$

$$P(A_1) = P\{O_1\, E_2\, O_3\, E_4\, E_5\} = \frac{5}{10}\cdot\frac{5}{9}\cdot\frac{4}{8}\cdot\frac{4}{7}\cdot\frac{3}{6} = \frac{5}{126}. \qquad [O = \text{odd},\ E = \text{even}]$$

$$\therefore \quad P(A) = 15/126 = 5/42.$$

3*. Four cards can be selected out of 52 cards in $\binom{52}{4}$ ways. To find the number of favourable cases, we exhibit the deck and possible favourable selection as under :

	Spade ace (1)	*Non-spade ace (3)*	*Spade non-ace (12)*	*Non-spade non-ace (36)*
A	1	1	1	1
B	0	2	2	0

There are two disjoint possibilities A and B as indicated above.

$$P(A) = \binom{1}{1}\binom{3}{1}\binom{12}{1}\binom{36}{1}\Big/\binom{52}{4} = \frac{1296}{270725};\quad P(B) = \binom{1}{0}\binom{3}{2}\binom{12}{2}\binom{36}{0}\Big/\binom{52}{4} = \frac{198}{270725}.$$

$$\therefore \quad p = P(A \uplus B) = (1296 + 198)/270725 = 1494/270725.$$

4*. Denote the drawing of "white ball from urn i by W_i and black ball from urn j and B_j". Now,

$$P(W_1 W_2 B_3) = \frac{1}{3}\cdot\frac{3}{4}\cdot\frac{3}{5} = \frac{9}{60}, \quad P(W_1 B_2 W_3) = \frac{1}{3}\cdot\frac{1}{4}\cdot\frac{2}{5} = \frac{2}{60}, \quad P(B_1 W_2 W_3) = \frac{2}{3}\cdot\frac{3}{4}\cdot\frac{2}{5} = \frac{12}{60}$$

$$\therefore \; p = P[W_1 W_2 B_3 \cup W_1 B_2 W_3 \cup B_1 W_2 W_3] = P(W_1 W_2 B_3) + P(W_1 B_2 W_3) + P(B_1 W_2 W_3)$$

$$= \frac{9}{60} + \frac{2}{60} + \frac{12}{60} = \frac{23}{60}.$$

5*. We are in for a chance when three balls are either all of one colour [event C_1] or these are of two colours [event C_2]. For event C_1, 3 red balls is the only obvious possibility.

$$C_2 = \{(2r, 1g) \uplus (2r, 1y) \uplus (2g, 1y) \uplus (2g, 1r)\}.$$

If L_1 denotes the event that at least one colour is not drawn, then $L_1 = C_1 \uplus C_2$ and

$$P(L_1) = P(C_1) + P(C_2)\left\{\binom{3}{3}\Big/\binom{6}{3}\right\} + \left\{\left[\binom{3}{2}\binom{2}{1} + \binom{3}{2}\binom{1}{1} + \binom{2}{2}\binom{1}{1} + \binom{2}{2}\binom{3}{1}\right]\Big/\binom{6}{3}\right\} = \frac{7}{10}.$$

6*. After the insertion of 3 yellow (y) balls, the composition of the box is as under : $A = (3w, 0b, 3y)$, $B = (2w, 1b, 3y)$, $C = (1w, 2b, 3y)$, $D = (0w, 3b, 3y)$. [w = white b = black] Let E = {three balls drawn out are of different colours}. Clearly, the compositions A and D are incompatible with E. We can then have; from sets B and C,

$$P(E) = \left\{\binom{2}{1}\binom{1}{1}\binom{3}{1}\Big/\binom{6}{3}\right\} + \left\{\binom{1}{1}\binom{2}{1}\binom{3}{1}\Big/\binom{6}{3}\right\} = 2\left[\frac{2.3}{20}\right] = \frac{3}{5}.$$

Note. $P_B(W_1 B_2 Y_3) = 2/6 \,.\, 1/5 \,.\, 3/4 = P_B(W_1 Y_2 B_3)$
$= P_B(B_1 W_2 Y_3) = P_B(B_1 Y_2 W_3) = P_B(Y_1 B_2 W_3) = P_B(Y_1 W_2 B_3).$

Thus, if *order* is considered even then $p = p_B + p_C = 6(1/20 + 1/20) = 3/5$.

7*. Let W, R, B denote the number of white, red and black balls among the four drawn out. Then

(a) $P\{B = 0\} = \binom{9}{4}\Big/\binom{15}{4} = \dfrac{126}{1365} = \dfrac{42}{455}.$

(b) $P\{B = 2, \bar{B} = 2\} = \binom{6}{2}\binom{9}{4}\Big/\binom{15}{4} = \dfrac{15 \times 36}{1365} = \dfrac{36}{91}.$

(c) $p = P\{(W = 4) \uplus (R = 4) \uplus (B = 4)\} = P(W = 4) + P(R = 4) + P(B = 4)$

$$p = \left[\binom{4}{4} + \binom{5}{4} + \binom{6}{4}\right]\binom{15}{4}^{-1} = (1 + 5 + 15)/1365 = 1/65.$$

(d) Let $E = \{(W = 1, R = 1, B = 2) \uplus (W = 1, R = 2, B = 1) \uplus (W = 2, R = 1, B = 1)\}$.

$\therefore \; P(E) = P(W = 1, R = 1, B = 2) + P(W = 1, R = 2, B = 1) + P(W = 2, R = 1, B = 1)$

$$= \frac{\binom{4}{1}\binom{5}{2}\binom{6}{1}}{1365} + \frac{\binom{4}{1}\binom{5}{1}\binom{6}{2}}{1365} + \frac{\binom{4}{2}\binom{5}{1}\binom{6}{1}}{1365} = (300 + 240 + 180)/1365 = 720/1365 = 48/91.$$

8*. We designate the events :

$C = \{\text{Void in Clubs}\}$, $D = \{\text{Void in Diamonds}\}$, $H = \{\text{Void in Hearts}\}$, $S = \{\text{Void in Spades}\}$.

Now, by symmetry,

$$P(C) = P(D) = P(H) = P(S);\ \ P(CD) = P(CH) = P(CS) = P(DH) = P(DS) = P(HS)$$

$P(CDH) = P(CDS) = P(DHS) = P(CHS);\ P(CDHS) = P(\varnothing) = 0$. [Void in all suits is impossible.]

So $p = P(C \cup D \cup H \cup S) = 4P(C) - 6P(CD) + 4P(CDH) - P(CDHS)$ [Poincare's formula]

$$p = 4\binom{39}{13}\binom{52}{13}^{-1} - 6\binom{26}{13}\binom{52}{13}^{-1} + 4\binom{52}{13}^{-1} = \left[4\binom{39}{13} - 6\binom{26}{13} + 4\right]\Big/\binom{52}{13}.$$

9*. We define the events as under :

$A = \{\text{All the four suits appear among the } n \text{ cards}\}$

$B = \{\text{At least one of the four suits does not occur among the } n \text{ cards}\}$

Obviously, $B = \overline{A}$ or $A = \overline{B}$ so that $P(A) = 1 - P(B)$ ...(1)

Let C, D, H, S denote the absence of clubs, hearts, diamonds and spades among the six cards, so

$$P(B) = P\{C \cup D \cup H \cup S\} = 4P(C) - 6P(CD) + 4P(CHD) - P(CDHS) \quad \text{...(2)}$$

Since, by symmetry, $P(C) = P(D)$, etc. Also $CDHS = \varnothing$, so $P(CDHS) = 0$. Because the cards are replaced, the probabilities do not alter from trial to trial. Thus, for n draws with replacement

$$P(C) = (3/4)^n, \quad P(CD) = (1/2)^n, \quad P(CDH) = (1/4)^n$$

Substituting into (2), using (1) we obtain

$$P(A) = 1 - 4(0.75)^n + 6(0.5)^n - 4(0.25)^n. \qquad [\text{If } n = 6,\ P(A) = 0.381]$$

10*. Let A, B, C denote the respective events that "at least one ball enters the box B_1, B_2, B_3 respectively". Now probability of a ball entering B_1 is 1/3, that of not entering B_1 is 2/3. Hence $P(\overline{A}) = (2/3)^n$, [none of the n balls enters B_1]. Similarly

$P(\overline{A}\,\overline{B}) = (1/3)^n$, since $\overline{A}\,\overline{B}$ means all the n balls enter B_3, and $P(\overline{A}\,\overline{B}\,\overline{C}) = P(\varnothing) = 0$.

$$P(\overline{A} \cup \overline{B} \cup \overline{C}) = 3P(\overline{A}) - 3P(\overline{A}\,\overline{B}) + P(\overline{A}\,\overline{B}\,\overline{C}) = 3(2/3)^n - 3(1/3)^n + 0 = (2^n - 1)/3^{n-1}.$$

If X denotes the number of empty boxes, then

(a) $P\{X = 0\} = 1 - P(\overline{A} \cup \overline{B} \cup \overline{C}) = (3^{n-1} - 2^n + 1)/3^{n-1}$.

(b) $P\{X = 2\} = P(\overline{A}\,\overline{B} \uplus \overline{B}\,\overline{C} \uplus \overline{C}\,\overline{A} \uplus \overline{A}\,\overline{B}\,\overline{C}) = 3P(\overline{A}\,\overline{B}) + P(\varnothing) = 1/3^{n-1}$.

(c) $P\{(X = 0) \uplus (X = 1) \uplus (X = 2\} = 1$.

$\therefore\ P(X = 1) = 1 - P(X = 0) - P(X = 2) = (2^n - 2)/3^{n-1}$.

11*. The following chain of subset relations is obvious :

$$\varnothing \subseteq A \cap B \subseteq A \subseteq A \cup B \subseteq A \uplus B. \quad \text{...(1)}$$

We apply Monotonicity of probability measure to (i), considering the pairs one at a time; this gives

$$P(\varnothing) \le P(AB) \le P(A) \le P(A \cup B) \le P(A \uplus B).$$

Since $P(\varnothing) = 0$, and $P(A \cup B) = P(A) + P(B) - P(AB) \le P(A) + P(B)$, we get

$$0 \le P(AB) \le \min\{P(A), P(B)\} \le P(A \cup B) \le P(A) + P(B) \qquad [AB \subseteq A, AB \subseteq B]$$

12*. From stated condition : $A \cap B \subset C$, we get $P(AB) \le P(C)$.

$P(A \cup B) = P(A) + P(B) - P(AB) \le 1, \Rightarrow P(A) + P(B) \le 1 + P(AB) \le 1 + P(C)$,

i.e., $P(C) \ge P(A) + P(B) - 1$.

By Negation $\quad 1 - P(\bar{C}) \ge 1 - P(\bar{A}) + 1 - P(\bar{B}) - 1 \quad \Rightarrow \quad P(\bar{C}) \le P(\bar{A}) + P(\bar{B})$.

13*. From Bon-Ferroni'a inequality $P(AB) \ge P(A) + P(B) - 1 = 0.9 + 0.8 - 1 = 0.7$.

14*. Let A denote the event that there is at least one pair among r shoes selected. Then A' denotes the event that there are no pairs among the r selected shoes. We investigate $P(A')$.

Observe that r shoes can be chosen out of $2n$ shoes in $N = \binom{2n}{r}$ ways. This is precisely the number of subsets of size r obtainable from a set of size $2n$.

Let us choose r pairs out of n pairs, this can be accomplished in $\binom{n}{r}$ ways. From each of these r pairs, we select only one of these two shoes. Obviously, there are $\binom{2}{1}\binom{2}{1}\cdots\binom{2}{1} = 2^r$ ways of doing it. By Sequential Counting, $f = \binom{n}{r}.(2^r)$ of the events in A' specify collection of shoes with no pair. Thus

$$P(A') = \binom{n}{r}\frac{2^r}{N}.\ P(A) = 1 - \left[\binom{n}{r}(2^r)\Big/\binom{2n}{r}\right].$$

L	L		L	L
R	R		R	R
A_1	A_2		A_{r-1}	A_r

Note. For $n = 10$, $r = 4$. $P(A) = 1 - \left(\dfrac{224}{323}\right) = \dfrac{99}{323}$.

15*. Let Q_k be the probability that in k drawings, none of the tickets bearing numbers 1, 2, 3, ..., n appear. Then $P_k = 1 - Q_k$ is the probability sought. We define the events : t_i = {ticket numbered i does not appear in one draw} ; T_i = {ticket numbered i does not appear in k draws}

$$\therefore \quad P(t_i) = \binom{n-1}{m}\Big/\binom{n}{m} = \frac{n-m}{n} = \left(1 - \frac{m}{n}\right)$$

$$P(T_i) = P\{(t_i)_1\,(t_i)_2 \ldots (t_i)_k\} = [P(t_i)]^k = [1 - (m/n)]^k$$

$$P(t_i \cap t_j) = \binom{n-2}{m}\Big/\binom{n}{m} = \left(1 - \frac{m}{n}\right)\left(1 - \frac{m}{n-1}\right). \quad P(T_i \cap T_j) = \left(1 - \frac{m}{n}\right)^k\left(1 - \frac{m}{n-1}\right)^k.$$

Similarly, $\qquad P(T_i \cap T_j \cap T_l) = \left(1 - \dfrac{m}{n}\right)^k\left(1 - \dfrac{m}{n-1}\right)^k\left(1 - \dfrac{m}{n-2}\right)^k, \ldots.$

Now, by Poincare's Formula : $Q_k = S_1 - S_2 + S_3 - S_4 + \ldots$

$$S_1 = \Sigma P(T_i) = \binom{n}{1}\left(1 - \frac{m}{n}\right)^k,\quad S_2 = \Sigma P(T_i T_j) = \binom{n}{2}\left(1 - \frac{m}{n}\right)^k\left(1 - \frac{m}{n-1}\right)^k,$$

$$S_3 = \Sigma P(T_i T_j T_e) = \binom{n}{3}\left(1-\frac{m}{n}\right)^k\left(1-\frac{m}{n-1}\right)^k\left(1-\frac{m}{n-2}\right)^k, \ldots$$

$$\therefore \quad P_k = 1 - Q_k = 1 - S_1 + S_2 - S_3 + S_4 - \ldots$$

$$= 1-\binom{n}{1}\left(1-\frac{m}{n}\right)^k+\binom{n}{2}\left(1-\frac{m}{n}\right)^k\left(1-\frac{m}{n-1}\right)^k-\binom{n}{3}\left(1-\frac{m}{n}\right)^k\left(1-\frac{m}{n-1}\right)^k\left(1-\frac{m}{n-2}\right)^k+\ldots$$

Sec. 1-63. Page 41

1*. From given data : $P(AB) = rp$, $P(B) = rp/q$, $P(A) = p$.

(a) $P(AB) = 0 \Rightarrow rp = 0$.

(b) $P(A) + P(B) = 1 \Rightarrow p(q + r) = q$; $rp = 0 \Rightarrow pq = q$.

(c) $A \subseteq B \Rightarrow A \cap B = A$, so $P(AB) = P(A) \Rightarrow rp = q$.

$B \subseteq A \Rightarrow AB = B$, so $P(AB) = P(B) \Rightarrow rp = (rp/q)$ or $rp(1 - q) = 0$.

(d) $0 = P(\overline{A}\overline{B}) = 1 - P(A \cup B) = 1 - [P(A) + P(B) - P(AB)] = 1 - p - (rp/q) + rp$.

Thus $\quad p(q + r) = q(1 + pr)$.

2*. Consider a biased tetrahedron with faces a, b, c, d with probabilities θ, 2θ, 2θ, 5θ ($\theta = 1$) of falling down. Let $A = \{a, b, d\}$, $B = \{b, c\}$, $C = \{a, b, c\}$, then $P(A) = 8\theta$, $P(B) = 4\theta$, $P(C) = 5\theta$. Obviously $P(A) > P(B)$. Also $P(AC) = 3\theta$, $P(BC) = 4\theta$. Clearly, $P(A/C) \not> P(B \mid C)$, [$\because$ $P(AC) < P(BC)$].

Note. If you take $B = C$, then $P(A) > P(B) \Rightarrow A$ has more atoms than B. As such, $P(A \mid C)$ which has a meaning if $A \subseteq C$ is ruled out.

3*. $P(C \mid AB) = P(C \mid B) \Rightarrow \dfrac{P(ABC)}{P(AB)} = \dfrac{P(BC)}{P(B)}$, i.e. $\dfrac{P(ABC)}{P(BC)} = \dfrac{P(AB)}{P(B)} \Rightarrow P(A \mid BC) = P(A \mid B)$

4*. From $A_1 \subseteq A_1 \cup A_2 \subset B_1B_2 \subset B_1$, it follows that $A_1 \subset B_1$. Similarly, $A_1 \subset B_2$, $A_2 \subset B_1$, $A_2 \subset B_2$.

Now $\quad P(A \mid B) = P(AB)/P(B) = P(A)/P(B) \quad [(\because A \cap B = A$, as $A \subset B]$ ...(1)

Using (1) we obtain

$$P(A_1 \mid B_1)\, P(A_2 \mid B_2) = \frac{P(A_1)}{P(B_1)} \cdot \frac{P(A_2)}{P(B_2)} = \frac{P(A_1)}{P(B_2)} \cdot \frac{P(A_2)}{P(B_1)} = P(A_1 \mid B_2)\, P(A_2 \mid B_1). \quad \text{[by (1)]}$$

5*. $\quad P(A \cap B) = P(A \cap \uplus_j B_j) = P(\uplus_j AB_j) = \Sigma P(AB_j) = \Sigma P(B_j)\, P(A \mid B_j)$ ...(i)

$= p\Sigma P(B_j) = p \,.\, P(B) \quad [\because P(B) = \Sigma P(B_j)]$ ...(ii)

$\therefore \quad P(A \mid B) = P(AB)/P(B) = pP(B)/P(B) = p$.

6*. Here $\quad \Omega = \{H, TH, TTH, TTT\}$; with the associated data :

$P(A) = 1/2, \quad P(TH) = 1/4, \quad P(TTH) = P(TTT) = 1/8.$

Let $\quad A = \{$No head on first toss$\} = \{TH, TTH, TTT\}$, so $P(A) = 1/2$.

Let $\quad B = \{$Coin is tossed three times$\} = (TTH, TTT\}$, so $P(B) = 1/4$.

$A \cap B = B$ so $P(AB) = P(B) = 1/4$. Consequently, $P(B \mid A) = P(AB)/P(A) = 2/4 = 1/2$.

7*. Let J_k denote that kth card is a Jack and T denote the occurrence of a ten. Now

(a) $P\{J_2 \mid J_1\} = \dfrac{P(J_1 J_2)}{P(J_1)} = \dfrac{\binom{4}{2}}{\binom{52}{2}} \Big/ \dfrac{\binom{4}{1}}{\binom{52}{1}} = \dfrac{1}{17}$. (b) $P(J_1 J_2) = \dfrac{\binom{4}{2}}{\binom{52}{2}} = \dfrac{1}{221}$.

(c) $$P(T \mid J_1) = \frac{P(J_1 T)}{P(J_1)} = \frac{4}{52}\cdot\frac{4}{51}\Big/\frac{4}{52} = \frac{4}{51}.$$

8*. We define the events :

A = {Exactly 3 defectives obtained in 9 examined}

B = {10th examined piece is defective}

Defective	Non-defective	
3	6	A
1	5	B

$\therefore\ P(A) = \binom{4}{3}\binom{11}{6}\Big/\binom{15}{9} = \frac{24}{65}$,

$P(B \mid A) = 1/6$, so $P(AB) = P(A)\,P(B \mid A) = 4/65$.

9*. Let B_i denote the event that a black ball is selected on the ith trial. We need $P(B_1 B_2 B_3)$. So

$$P(B_1 B_2 B_3) = P(B_1)\,P(B_2 \mid B_1)\,P(B_3 \mid B_1 B_2) = \frac{3}{10}\cdot\frac{5}{12}\cdot\frac{7}{14} = \frac{1}{16}.$$

10*. Let the four cards drawn be A, B, C, D, then

$P(ABCD) = P(A)\,.\,P(B \mid A)\,.\,P(C \mid AB)\,.\,P(D \mid ABC)$ [Product Rule] ...(1)

(a) Now $P(A) = 40/40$, $P(B \mid A) = 30/39$, $P(C \mid AB) = 20/38$, $P(D \mid ABC) = 10/37$.

$$\therefore\quad P(ABCD) = \frac{40}{40}\frac{30}{39}\frac{20}{38}\frac{10}{37} = \frac{1000}{9139}.$$

(b) Here $P(A) = 40/40$; cancel suit and denomination of A to get $P(B \mid A) = 27/39$. Similarly,

$$P(C \mid AB) = 16/38, \qquad P(D \mid ABC) = 7/37$$

$$\therefore\quad P(ABCD) = \frac{40}{40}\cdot\frac{27}{39}\cdot\frac{16}{38}\cdot\frac{7}{37} = \frac{504}{9139}.$$

11*. Let A = {At least one 6 in 4 tosses of a single die}, B = {At least one double-6 in 24 tosses of 2 fair dice}.

Then $P(A) = 1 - P(\bar{A}) = 1 - P(\text{no six in 4 tosses of one die}) = 1 - (5/6)^4 = 0.518$.

$P(B) = 1 - P(\bar{B}) = 1 - P(\text{no double-6 in 24 tosses of dice-pair}) = 1 - (35/36)^{24} = 0.491$.

12*. (i) Let B_j denote that jth ball withdrawn, is not black. Then

$$P(\bar{B}_1) = 47/50, \quad P(\bar{B}_2) = 46/49, \quad P(\bar{B}_3) = 45/48,$$

$$P(\bar{B}_{r-1}) = \frac{47-(r-2)}{50-(r-2)},\ P(B_r) = \frac{3}{50-(r-1)}$$ [rth : black ball].

Thus, the probability P_r, the rth ball drawn is the first black ball, is

$$P_r = \frac{47}{50}\cdot\frac{46}{49}\frac{45}{48}\frac{44}{47}\frac{43}{46}\cdot\frac{42}{45}\cdots\frac{47-(r-2)}{50-(r-2)}\cdot\frac{3}{50-(r-1)}$$ [Product Rule]

$$= \frac{3(49-r)(50-r)(51-r)}{50.49.48(51-r)} = \frac{(49-r)(50-r)}{39200}.$$

(ii) In case balls are replaced

$$P(\bar{B}_1) = 47/50 = P(\bar{B}_2) = \ldots = P(\bar{B}_{r-1});\ P(B_r) = 3/50$$

$$\therefore\ P_r = \left(\frac{47}{50}\frac{47}{50}\cdots\frac{47}{50}\right)\cdot\frac{3}{50} = \left(\frac{47}{50}\right)^{r-1}\frac{3}{50} = \frac{3\times(47)^{r-1}}{(50)^r}.$$

13*. Let A, B, C, D denote the events that four players get a complete suit. By symmetry of distribution

$P(A) = P(B) = P(C) = P(D)$, $P(AB) = \ldots = P(CD)$, etc. Now by Poincare's Formula

$$p = P(A \cup B \cup C \cup D) = S_1 - S_2 + S_3 - S_4$$
$$= 4P(A) - 6P(AB) + 4P(ABC) - P(ABCD) \qquad [P(ABCD) \equiv P(ABC)] \quad \ldots(1)$$

As $P(A) = 4/{}^{52}C_{13}$, $P(AB) = P(A)\, P(B \mid A) = \left\{4\Big/\binom{52}{13}\right\}\left\{3\Big/\binom{39}{13}\right\}$

$$P(ABC) = P(C)\, P(C \mid AB) = \left\{4\Big/\binom{52}{13}\right\}\left\{3\Big/\binom{39}{13}\right\}\left\{2\Big/\binom{26}{13}\right\}$$

Substitutions into (1) provide, after simplification

$$p = \{16(39!\ 13\ !) - 72[13\ !\ 13\ !\ 26\ !] + 72(13!)^4\}/52!.$$

14*. Definition : Two events A and B are said to be equivalent : A occurs iff B occurs, i.e., they are equal as sets. Thus, if A and B are equivalent, they have the same probability measure, i.e., $P(A) = P(B)$.

$$P(A \cup B \cup C) = \sum P(A) - \sum P(AB) + P(ABC) \qquad \text{[Poincare's Formula]}$$

Using given condition this simplifies to $\sum P(A) - \sum P(AB) = 0$.

$\therefore$ $[P(A) - P(AB)] + [P(B) - P(BC)] + [P(C) - P(CA)] = 0$.

or $P(A)\,[1 - P(B \mid A)] + P(B)\,[1 - P(C \mid B)] + P(C)\,[1 - P(A \mid C)] = 0$.

This result can hold iff $P(B \mid A) = P(C \mid B) = P(A \mid C) = 1$. ...(i)

Similarly $P(A \mid B) = P(B \mid C) = P(C \mid A) = 1$. ...(ii)

From (i) and (ii) we conclude that $P(A) = P(B) = P(C)$. Consequently A, B, C are equivalent.

Sec. 1-90. Page 52

1*. Since $A \cap B = \varnothing$, we must have $A \subseteq \bar{B}$ and $B \subseteq \bar{A}$. [Draw Figure]

Hence by Monotonic law, $P(A) \le P(\bar{B})$ and $P(B) \le P(\bar{A})$.

Aliter : $P(A \uplus B) \le 1 \Rightarrow P(A) + P(B) \le 1$. ...(i)

Thus $P(A) \le P(\bar{B})$ and $P(B) \le P(\bar{A})$. [both by Complement Rule]

2*. Denote the events : C = {both balls are of the same colour}, N = {both balls bear the same number}. Since the two balls of the same colour have the same number, $C \cap N = \varnothing$. Thus

$$P(C \uplus N) = P(C) + P(N). \qquad \ldots(1)$$

The number of ways in which the pair (two balls) consists of 2 white balls is $\binom{n}{2}$; same is true for the pair to consist of two red balls or two black balls. Hence by the Rule of Sum, the total number of ways in which the pair chosen consists of two balls of the same colour is $3\binom{n}{2}$ ways. Since the total number of ways in which 2 balls can be chosen out of $3n$ balls is $\binom{3n}{2}$; hence $P(C) = 3\binom{n}{2}\binom{3n}{2}$.

Both chosen balls bear No. 1 if these are white-red, white-black or red-black. Thus, for cardinality 1, there are 3 ways. Same is true for numbers 2 through n. Thus, there are $3n$ ways for N to materialize. Consequently, $P(N) = 3n\Big/\binom{3n}{2}$.

$$P(C \uplus N) = P(C) + P(N) = \left\{3\binom{n}{2}\Big/\binom{3n}{2}\right\} + \left\{3n\Big/\binom{3n}{2}\right\} = \frac{n+1}{3n-1}.$$

3*. Let A, B, C denote the events that the misplacement happened in Iowa, in New York and in London, respectively and M denote the event that luggage has been misplaced. Then A, B, C are incompatible (m-e) events and $M = A \uplus \bar{A}B \uplus \bar{A}\bar{B}C$.

$P(M) = P(A) + P(\bar{A}B) + P(\bar{A}\bar{B}C) = P(A) + P(B).P(\bar{A}) + P(\bar{A}).P(\bar{B}).P(C)$

where we believe that staff in different airports misbehave independently of one another. Thus, using $p + q = 1$, the needed probabilities are :

$P(M) = q + pq + p^2q = q(1 + p + p^2) = q(1 - p^3)/(1 - p) = 1 - p^3.$

$P(A \mid M) = P(AM)/P(M) = P(A)/P(M) = q/(1 - p^3).$

$P(\bar{A}B \mid M) = P(\bar{A}B \cap M)/P(M) = P(\bar{A}B)/P(M) = qp/1(1 - p^3).$

$P(\bar{A}\bar{B}C \mid M) = P(\bar{A}\bar{B}C \cap M)/P(M) = P(\bar{A}\bar{B}C)/P(M) = p^2q/(1 - p^3).$

Chapter 2 : Probabilistic Independence. Baye's Reversal Rule

Sec. 2-12. Page 63

1*. Write $P(B) = x$. By Multistage Rule

$$P(A) = P(B)P(A \mid B) + P(\bar{B})P(A \mid \bar{B}) = xP(A \mid B) + (1 - x)P(A \mid \bar{B}).$$

By Section Formula of Geometry, this result implies that $P(A)$ lies between $P(A \mid B)$ and $P(A \mid \bar{B})$.

The extreme values of $P(A)$ are attained when $x = 0$ and $x = 1$. Thus, the two possibilities are

$$P(A \mid B) \le P(A) \le P(A \mid \bar{B})\ ;\quad P(A \mid \bar{B}) \le P(A) \le P(A \mid B). \qquad ...(1)$$

When A and B are independent, each member of (1) is merely $P(A)$.

2*. Define the events : E = {2nd card C_2 covers the first card C_1}

S = {Cards C_1 and C_2 drawn out are of the **same** suit}

D = {Cards C_1 and C_2 drawn out are of the **different** suits}

$\therefore$ $P(E \mid D) = 0$ (trivally). $P(S) = \frac{52}{52}\cdot\frac{12}{51} = \frac{4}{17}$ [C_1 is of arbitrary suit and C_2 has same suit as that of C_1]

In case event S occurs, the equiprobability provides

$P\{C_2 \text{ covers } C_1\} = P\{C_1 \text{ covers } C_2\} \Rightarrow P\{E \mid S\} = 12$

$\therefore$ $P(E) = P(S)\,P(E \mid S) + P(D)\,P(E \mid D) = \frac{4}{17} \times \frac{1}{2} + 0 = \frac{2}{17}.$ [Multistage p-Rule]

Note. **Ace** is regarded superior to **King**. We may assign values 2, 3, ..., .13, 14 (ace) to the cards in any other method of evaluation.

3*. The sample space consists of three m.e. possibilities (events) C_1, C_2, C_3 as under : C_1 = {A_s is the 1st card (hence K_s is 2nd card)}; C_2 = {A_s is the last 152nd) card (hence K_s is 51st card)].

$C_3 = \{A_s$ is somewhere in the middle ($K_s\,A_s$ or $A_s\,k_s$ occur)$\}$. Let $E = \{K_s$ is next to $A_s\}$
$S = C_1 \uplus C_2 \uplus C_3$; $P(C_1) = P(C_2) = 51\,!/52\,! = 1/52$, $P(C_3) = 50C_1/52C_1 = 50/52$
Observe that only one card is next to the top or bottom cards and 2 cards are next to a card within the deck, we have $P(E \mid C_1) = 1/51 = P(E + C_2)$.

$P(E \mid C_3) = 2/51$, $[K_s A_s \uplus A_s\, K_s]$. Now by, Multi-stage p-Rule.

$$P(E) = \Sigma P(C_1)\, P(E \mid C_1) = \frac{1}{52}\cdot\frac{1}{51} + \frac{1}{52}\cdot\frac{1}{51} + \frac{50}{52}\cdot\frac{2}{51} = \frac{2}{52} = \frac{1}{26}.$$

4*. Define the events :

E = {person selected is employed}, M = {person chosen is man}, W = {person chosen is women}. Then $P(M) = P(W) = \frac{1}{2}$. Now by Multi-stage p-Rule.

$$P(E) = P(M)\, P(E \mid M) + P(W)\, P(E \mid W) = \frac{1}{2}\cdot\frac{90}{100} + \frac{1}{2}\cdot\frac{55}{200} = 0.725.$$

5*. Let the urns be A and B, and E be the event of interest. Then by Multi-stage p-Rule, (since $P(A) = P(B) = 1/2$).

$$P(E) = P(A)\, P(E \mid A) + P(B)\, P(E \mid B) = \tfrac{1}{2}[P(E \mid A) + P(E \mid B)]$$

$E = \{3 \text{ or } 6\}$, $P\{3 \text{ or } 6\} = \frac{1}{2}\left[\frac{2}{6} + \frac{1}{5}\right] = \frac{4}{5}$.

$$P\{4, 6\} = \frac{1}{2}\left[\frac{2}{6} + \frac{2}{5}\right] = \frac{11}{30},\ P\{1, 10\} = \frac{1}{2}\left[\frac{1}{6} + \frac{1}{5}\right] = \frac{11}{60},\ P\{5\} = \frac{1}{2}\left[\frac{1}{6} + 0\right] = \frac{1}{12},\ P\{10\} = \frac{1}{2}\left[0 + \frac{1}{5}\right] = \frac{1}{10}.$$

6*. Let R_1, B_1, R_2 denote the events that the first ball drawn is red, first ball drawn is black and the second ball drawn is red. Then

$P\{R_1\} = \frac{3}{9}$, $P\{B_1\} = \frac{6}{9}$, $P\{R_2 \mid R_1\} = \frac{2}{8}$, $P(R_2 \mid B_1\} = \frac{3}{8}$

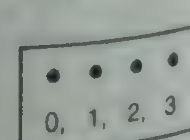
$\therefore P\{R_2\} = P\{R_1\}\, P(R_2 \mid R_1\} + P\{B_1\}\, P\{R_2 \mid B_1\} = \frac{3}{9}\cdot\frac{2}{8} + \frac{6}{9}\cdot\frac{3}{8} = \frac{24}{72} = \frac{1}{3}$. [Multi-stage p-Rule]

Thus result is, as if nothing is **thrown away** and we draw a red ball as is from original contents

Note. If we know the colour of the first ball, probability $P\{R_2\}$ is changed.
(i) $p = 2/8$, if the first ball is red. (ii) $p = 3/8$, if the first ball is back.

7*. Let B denote the rth ball and $B_r = k$ means rth ball chosen is numbered k. By Multi-stage p-Rule

$$P\{B_2 = 3\} = \Sigma P(B_1 = r)\, P\{(B_2 = 2) \mid (B_1 = r)\} \quad 0 \le r \le 3 \qquad 0, 1, 2, 3$$
$$= \frac{1}{4}\left[\frac{1}{3} + \frac{1}{3} + \frac{1}{2} + 0\right] = \frac{7}{23}, \quad \left[P(B_1 = r) = \frac{1}{4}\right]$$



8*. The probability that the first sample S_1 contains t white ball is

$$p_1 = \binom{W}{t}\binom{N-W}{n-t}\Big/\binom{N}{n}.$$

The probability that the second sample S_2 of size m contains k white balls is

$$p_2 = \binom{t}{k}\binom{n-t}{m-k}\Big/\binom{n}{m}.$$

We now use Theorem of Total Causes [i.e., Multi-Stage Rule] and note that $W > n$ and so get $p = P\{S_2 \text{ contains } k \text{ white balls}\}$

$$= \sum_{i=j}^{W} \frac{\binom{W}{t}\binom{N-W}{n-t}\binom{t}{k}\binom{n-t}{m-k}}{\binom{N}{n}\binom{n}{m}} = \binom{W}{k}\binom{N-W}{m-k}\Big/\binom{N}{m}. \text{ by reduction.}$$

Observe that, the first sample has no bearing on that final outcome.

9*. Let C_n be the event that a family has exactly n children and B_0 be the event that a family has no boy. Then by hypothesis, $P\{C_n\} = (1/2)^n$ and given C_n, B_0 requires n girls, we must have $P\{B_0 \mid C_n\} = (1/2)^n$, for every $n = 1, 2, \ldots$ By Multi-stage Rule :

$$P(B_0) = \sum_{n=1}^{\infty} P(C_n)\, P(B_0 \mid C_n) = \sum_{n=1}^{\infty} (\tfrac{1}{2})^{2n} = \tfrac{1}{4} \sum_{n=0}^{\infty} (\tfrac{1}{2})^n = \tfrac{1}{3}.$$

Note that : $P\{C_n \mid B_0\} = P\{B_0 \mid C_n\}\, P(C_n)/P(B_0) = 3/4$.

10*. We define the events : A = {First a tosses alł result in heads}

B = {Last b tosses all result in heads}, C_k = {Coin chosen has : $P(H) = k/n$}

We require $P(B \mid A)$, which by Multi-stage Rule, is given by

$$P(B \mid A) = \frac{P(AB)}{P(A)} = \frac{\Sigma\, P(AB \mid C_k)\, P(C_k)}{\Sigma\, P(A \mid C_k)\, P(C_k)} = \frac{\Sigma (k/n)^{a+b}\,(1/n)}{\Sigma (k/n)^a\,(1/n)}, 1 \le k \le n \qquad \ldots(1)$$

because, by the assumption concerning the method of choosing the coin to be tossed, we have, *a priori*, that $P(C_k) = 1/n$. Now using Euler-Maclaurin's summation formula,

$$\lim_{n\to\infty} \sum_{k=1}^{n} \left(\frac{k}{n}\right)^a \left(\frac{1}{n}\right) = \int_0^1 x^a\, dx = \frac{1}{a+1}. \qquad \ldots(2)$$

Hence, when n is large, using (2) into (1) we get,

$$P(B \mid A) = \frac{1/(a+b+1)}{1/(a+1)} = \frac{a+1}{a+b+1}, \qquad (n \text{ large}).$$

11*. We define the following events :

C_m= {A family has m children}, B_m = {A family has m boys}

We are given : $P_m = P\{C_m\} = \alpha p^m$, $m = 0, 1, 2, 3, \ldots$ $\qquad \ldots(1)$

Since each sex distribution is equiprobable,

$$P\{B_k \mid C_m\} = \binom{m}{k}\left(\tfrac{1}{2}\right)^{m-k}\left(\tfrac{1}{2}\right)^k = \binom{m}{k}\left(\tfrac{1}{2}\right)^m, m \ge k \qquad \ldots(2)$$

$$P\{B_k\} = \sum_{m=k}^{\infty} P(C_m)\, P\{B_k \mid C_m\} = \Sigma \alpha p^m . \binom{m}{k}\left(\frac{1}{2}\right)^m, \text{[by (1) and (2)] [Multi-stage Rule]}$$

$$= \alpha \sum_{r=0}^{\infty} \binom{k+r}{r}\left(\frac{p}{2}\right)^{k+r} = \alpha\left(\frac{1}{2}p\right)^k \left(1 - \frac{1}{2}p\right)^{-(k+1)}, \qquad \text{[Using } m-k=r\text{]}$$

$$= 2\alpha p^k/(2-p)^{k+1}. \qquad \left[\because (1-T)^{-n} = \sum_{r=0}^{\infty}\binom{n+r-1}{r} T^r\right]$$

(b) Let L_1 and L_2 denote the events that the family has at least one boy and at least two boys. Then,

$$P(Y \geq 1) = P\{L_1\} = \sum_{k=1}^{\infty} P(B_k) = \frac{2\alpha}{2-p} \sum_{k=1}^{\infty} \left(\frac{p}{2-p}\right)^k = \frac{2\alpha}{2-p} \cdot \frac{p/(2-p)}{1-[p/(2-p)]} = \frac{\alpha p}{(1-p)(2-p)}$$

$$P(Y \geq 2) = P\{L_2\} = \sum_{k=2}^{\infty} P(B_k) = \frac{2\alpha}{2-p} \sum_{k=2}^{\infty} \left(\frac{p}{2-p}\right)^k = \frac{2\alpha}{2-p} \cdot \frac{p^2/(2-p)^2}{1-[p/(2-p)]} = \frac{\alpha p^2}{(1-p)(2-p)^2}$$

Observe that $P\{Y \geq 1, Y \geq 2\} = P\{Y \geq 2\}$, hence

$$P\{Y_2 \geq 1 \mid Y_1 \geq 1\} = \frac{P(Y_2 \geq 1, Y_1 \geq 1)}{P(Y_1 \geq 1)} = \frac{P(Y_2 \geq 1)}{P(Y_1 \geq 1)} = \frac{p}{2-p}.$$

Sec. 2.25. Page 75

1*. $P(B) = 1 - P(\overline{B}) = \frac{1}{2}$, $P(A \cup B) = P(A) + P(B) - P(AB) \Rightarrow \frac{5}{6} = P(A) + \frac{1}{2} - \frac{1}{3} \Rightarrow P(A) = \frac{2}{3}$.

As $P(AB) = \frac{1}{3} = P(A) \,.\, P(B)$, it follows that A and B are independent.

2*. Suppose on the contrary, that the events are independent.

$$\therefore \quad P(A_1A_2A_3) = P(A_1)\,P(A_2)\,P(A_3) = p_1\,p_2\,p_3. \qquad \ldots(1)$$

The stated property then provides $P(A_1A_2) = P(A_2A_3) = P(A_3A_1 = P(A_1A_2A_3) \Rightarrow p_1p_2 = p_2p_3$ $= p_3p_1 = p_1p_2p_3$.

These relations imply that $p_1 = 1$, $p_2 = 1$, $p_3 = 1$ which contradicts the hypothesis $0 < p_i < 1$. Hence, these events are not independent.

Concrete Illustration. We throw two dice D_1 and D_2 and define the events :

$A = \{D_1 \text{ shows a } 6\} = \{(6, 1), (6, 2), (6, 3), (6, 4), (6, 5), (6, 6)\}$; $P(A) = 1/6$

$B = \{D_2 \text{ shows a } 6\} = \{(1, 6), (2, 6), (3, 6), (4, 6), (5, 6), (6, 6)\}$; $P(B) = 1/6$

$C = \{D_1, D_2 \text{ show the same number}\} = \{(1, 1), (2, 2), \ldots, (6, 6)\}$; $P(C) = 1/6$

$A \cap B = B \cap C = C \cap A = A \cap B \cap C = \{(6, 6)\} \Rightarrow P(AB) = P(BC) = P(CA) = P(ABC)$.

Obviously, $P(AB) = 1/36 = P(A)\,P(B)$, etc. But $P(ABC) = 1/36 \neq P(A)\,P(B)\,P(C) = 1/216$.

Consequently, these events are not independent.

3*. Firstly we not that if $P(B) = 1$, then $I(A\,;\,B)$ and $I(B\,;\,C)$ are always true. To see this, we notice that $A \cup B \supseteq B \Rightarrow P(A \cup B) \geq P(B) = 1$.

Since no probability measure can exceed unity, this result gives

$$P(A \cup B) = 1 \Rightarrow P(A) + P(B) - P(AB) = 1.$$

Thus, $\quad P(AB) = P(A) = P(A)\,.\,1 = P(A)\,P(B)$, so that $I(A\,;\,B)$ holds.

We now choose A and C to be dependent events; the consequences are $I(A\,;\,B)$ and $I(B\,;\,C)$ but these relation do not imply $I(A\,;\,C)$ because by hypothesis A and C are dependent events.

Note. This example shows that $I(A\,;\,B)$ is not an *equivalence* relation.

4*. $P[(AB) \cap C] = P(ABC) = P(A)\,P(B)\,P(C) = P(AB)\,P(C)$ since, $P(AB) = P(A)\,P(B)$.

This show that $I(AB\,;\,C)$ is true.

Now let $B = \varnothing$ and suppose A and C are dependent events. Then

$$P(AB) = P(\varnothing) = 0 = P(A)\,P(B)\,;\; P(ABC) = P(\varnothing) = 0 = P(A)\,P(B)\,P(C).$$

It follows that, the events A and C are dependent though the above conditions hold

5*. (i) This is a false assertion. Roll two dice D_1 and D_2 and let $A = \{\text{sum of eyes is } 7\}$.

$$B = \{D_1 \text{ does not show face } 6\},\quad C = \{D_2 \text{ does not show face } 1\}$$

Then $$P(A \mid B \cup C\} = \frac{P\{7, \text{not } (6,1)\}}{P\{\text{not } (6,1)\}} = \frac{5/36}{35/36} = \frac{1}{7} \neq P(A).$$

(ii) $P\{A(B \uplus C)\} = P\{AB \uplus AC\} = P(AB) + P(AC) = P(A)\,[P(B) + P(C)] = P(A).\, P(B \uplus C)$. [True]

6*. By Poincares' formula and independence

$$P(AB \cup BC \cup CA) = P(AB) + P(BC) + P(CA) - 3P(ABC) + P(ABC)$$
$$= P(A)\,P(B) + P(B)\,P(C) + P(C)\,P(A) - 2P(A)\,P(B)\,P(C) = pq + qr + rp - 2pqr.$$

7*. Let $A \cup B = C$, $A \cap B = D$, then using $P(AB) = P(A)\,P(B)$,

$P(C) = P(A) + P(B) - P(A)\,P(B)\,;\; P(D) = P(A)\,P(B),\; P(C \cap D) = P(D) \neq P(C)\,P(D)$

$[\because C \cap D = D]$

This establishes that C and D are *not* indep. events.

8*. $P\{A \cap (B \cup C)\} = P\{AB \cup AC\} = P(AB) + P(AC) - [P(A)\,(BC)]$ ...(i)

$= P(A)\,P(B) + P(A)\,P(C) - P(A)\,P(BC) = P(A)\,[P(B) + P(C) - P(BC)]$

$= P(A)\,P(B \cup C)$. [i.e., A is independent of $B \cup C$]

Remark. A is *Completely independent* of B and C means A is independent of B, C and $B \cap C$. That is why in (i), $P(ABC) = P(A)\,P(BC)$.

9*. (a) $P(AB) = P(A)\,P(B \mid A) = \frac{1}{8}$, $P(B) = P(AB)/P(A \mid B) = \frac{1}{2}$.

(i) Since $P(AB) = \frac{1}{8} \neq 0$. $A \cap B \neq \varnothing$. [False]

(ii) Since $A \subseteq B \Rightarrow AB = A$, so $P(AB) = P(A)$. This is false too.

(iii) Since $P(A) = P(A \mid B) \Rightarrow I(A\,;\,B) \Rightarrow I(\bar{A}\,;\bar{B})$ hence

$$P(\bar{A} \mid \bar{B}) = P(\bar{A}) = 1 - P(A) = \tfrac{3}{4} \qquad \text{[True]}$$

(iv) Since A and B are independent, it follows that

$$P(A \mid B) + P(A \mid \bar{B}) = P(A) + P(A) = 2P(A) = \tfrac{1}{2} \neq 1. \qquad \text{[False]}$$

(b) $$P(C \mid AB) = \frac{P(ABC)}{P(AB)} = \frac{P(A)\,P(BC)}{P(A)\,P(B)} \qquad \text{[By } I(A\,;\,B) \text{ and } I(A\,;\,BC)]$$

$$= \frac{P(A)\,.\,P(B)\,.\,P(C)}{P(A)\,.\,P(B)} = P(C). \qquad \text{[by } I(B\,;\,C)]$$

10*. Let $P(B) = y$ and $P(C) = z$, set $1 - a = a'$, $1 - b = b'$, $1 - c = c'$, $1 - y = y'$, $1 - z = z'$,

Now $x = P(A'B'C) = P\{(A \cup B)'\,C] = P(C) - P[(A \cup B)\,C] = P(C) - P[(AC \cup BC)]$

$= P(C) - P(AC) - P(BC) + P(ABC) = P(C) - P(A)\,P(C) - P(B)\,P(C) + P(A)\,P(B)\,P(C)$, [by Indep.]

$= z\{1 - a - y + ay\} = z(1 - a)\,(1 - y) = za'\,y'$ ...(1)

$b = 1 - P(A \cup B \cup C) = P(A'B'C') = P(A')\,P(B')\,P(C') = a' \,.\, y' \,.\, z'$ ...(2)

$c = 1 - P(ABC) = 1 - P(A)\,P(B)\,P(C) = 1 - ayz$ ...(3)

From (1) and (2) : $x/b = z/z' \Rightarrow z = x/(x + b) = P(C)$. ...(4)

From (4) and (3) : $\dfrac{axy}{x+b} = c' \;\Rightarrow\; y = \dfrac{(x+b)\,c'}{ax} = P(B)$. ...(5)

From (1), (4) and (5) : $x = \left(\dfrac{x}{x+b}\right) a' \left(\dfrac{x(a-c') - bc'}{ax}\right)$ $\left[y' = 1 - y = \dfrac{x(a-c') - bc'}{ax}\right]$

$\therefore\; ax\,(x + b) = a'\,[x(a - c') - bc']$, or $ax^2 + [ab + a'\,(c' - a)]\,x + a'\,bc' = 0$. ...(6)

This is equivalent to Eq. (*A*). Further, x being *a* probability value, both roots x_1, x_2 of Eq. (6) must be positive and hence $x_1 + x > 0$. This yields that the Coefficient of x be negative. Thus

$$a'(a - c') > ab \;\Rightarrow\; a - \frac{ab}{a'} > 1 - c \;\Rightarrow\; c > a' + \frac{ab}{a'} = \frac{(1-a)^2 + ab}{(1-a)}.$$

11*. Let $P(A)$ mean the probability that A can solve the problem, with similar meanings attached to $P(B)$ and $P(C)$. Since A, B, C solve problem independently, the required probability p is

$$p = P\{A \cup B \cup C\} = 1 - P(\bar{A})\,P(\bar{B})\,P(\bar{C}) = 1 - \tfrac{1}{2}\cdot\tfrac{2}{3}\cdot\tfrac{3}{4} = \tfrac{3}{4}.$$

12*. Here $P(A) = p_1 = \frac{3}{5}$, $P(B) = p_2 = \frac{2}{5}$, $P(C) = p_3 = \frac{3}{4}$.

$P(ABC) = p_1p_2p_3 = \frac{3}{5}\cdot\frac{2}{5}\cdot\frac{3}{4} = \frac{18}{100}$ [all three hit the target independently]

$P(\bar{A}BC) = q_1p_2p_3 = \dfrac{12}{100}$, $P(A\bar{B}C) = p_1q_2p_3 = \dfrac{27}{100}$ $P(AB\bar{C}) = p_1p_2q_3 = \dfrac{6}{100}$ (C alone misses).

Let us define the events and their probabilities :

$P(E_1) = P\{\text{Exactly one of the shots hit}\} = P\{\bar{A}\bar{B}C \cup \bar{A}B\bar{C} \cup A\bar{B}\bar{C}\} = q_1q_2p_3 + q_1p_2q_3 + p_1q_2q_3.$

$P(E_2) = P\{\text{Exactly 2 shots hit}\} = P\{\bar{A}BC \cup AB\bar{C} \cup AB\bar{C}\} = q_1p_2p_3 + p_1q_2p_3 + p_1p_2q_3.$

$P(L_2) = P\{\text{At least 2 shots hit}\} = P\{E_2 \uplus ABC\} = P(E_2) + p_1p_2p_3.$

(a) $P(E_2) = (12 + 27 + 6)/100 = 45/100$.

(b) $P(E_1) = (18 + 4 + 9)/100 = 31/100$.

(c) $P\{L_2\} = P(E_2) + P(ABC) = 0.45 + 0.18 = 0.63$.

(d) $P\{A \mid E_1\} = P(A \cap E_1)/P(E_1) = P(\bar{B}\,\bar{C}A)/P(E_1) = (9/100)/(31/100) = 9/31$.

(e) $P\{\bar{C} \mid E_2\} = P\{AB\bar{C}\}/P(E_2) = 6/45 = 2/15$.

13*. Since A_i are independent events so $\bar{A}_i$ must also be independent events. Thus,

$$p = P\{\bar{A}_1, \bar{A}_2, \ldots, \bar{A}_n\} = P(\bar{A}_1)\,P(\bar{A}_2) \ldots P(\bar{A}_n) = (1 - p_1)(1 - p_2) \ldots (1 - p_n) \quad \ldots(1)$$

To find the stated inequality, we observe that

$$e^{-t} \geq (1 - t),\; 0 \leq t < 1. \quad \ldots(2)$$

Replacing t by $p_1, p_2, \ldots, p_n$, we get by repeated application of (2), using (1) :

$$(1-p_1)(1-p_2)\ldots(1-p_n) \le e^{-p_1} . e^{-p_2} \ldots e^{-p_n} \Rightarrow p \le \exp(-\Sigma p_i).$$

Note. If $\Sigma p_i \to \infty$ then $p \to 0$ as $n \to \infty$.

14*. A figure for three events A_1, A_2, A_3 instantly provides :

$$z \notin A_1 A_2 A_3 \Rightarrow a \notin (\bar{A}_1 \bar{A}_2 \bar{A}_3 \uplus A_1 \bar{A}_2 \bar{A}_3 \uplus \bar{A}_1 A_2 \bar{A}_3 \uplus \bar{A}_1 \bar{A}_2 A_3).$$

$\therefore P\{z \notin A_1 A_2 A_3\} = q_1 q_2 q_3 + p_1 q_2 q_3 + q_1 p_2 q_3 + q_1 q_2 p_3 = q[1 + (p_1/q_1) + (p_2/q_2) + (p_3/q_3)], [q = q_1 q_2 q_3]$

We now extend it to r subsets. Let $A = A_1 A_2 \ldots, A_r$; then $(x_1 \notin A), x_2 \notin A, \ldots, x_N \notin A$

$p = P\{x_1 \notin A, \ldots, x_N \notin A\} = (P\{Px_1 \notin A\})^N = \{q[1 + \Sigma(p_i/q_i)\}^N, 1 \le i \le r, q = q_1 q_2 \ldots q_r.$

Sec. 2-53. Page 87

1*. Let A = {sum total of 9 on three dice}. A suitable decomposition of S is as under :
$B_1 = \{1, 2, 6\}, B_2 = \{1, 3, 5\}, B_3 = \{1, 4, 4\}, B_4 = \{2, 2, 5\}, B_5 = \{2, 3, 4\}, B_6 = \{3, 3, 3\}$.
B_7 = {sum total different from 9}. Various triplets are in some order.
Obviously, $P(AB_j) = P(B_j), j = 1, 2, \ldots, 6$ $P(AB_7) = 0$. By Baye's Reversal Rule :
$P(B_6 \mid A) = P(B_6)/[P(B_1) + \ldots P(B_6)] = p/(6p + 6p + 3p + 3p + 6p + p) = 1/25$. $[p = 1/216]$.

2*. Since $P(H) = P(C_1)\, P(H \mid C_1) + P(C_2)\, P(H \mid C_2) = (3/4)(1/2) + (1/4) . 1 = 5/8$

$\therefore\ P(C_2 \mid H) = P(C_2) . P(H \mid C_2)/P(H) = (1/4) . 1/(5/8) = 2/5.$

3*. If $P(A)$ denotes the probability that A is appointed manager, with like meanings for $P(B)$ and $P(C)$, then we are given that $P(A) = 4/9$, $P(B) = 2/9$, $P(C) = 3/9$.

If E denotes the event : Bonus Scheme is introduced; then we have

$$P(E \mid A) = 0.3,\ P(E \mid B) = 0.5,\ P(E \mid C) = 0.8$$

$$\therefore\ P(E) = P(A)\, P(E \mid A) + P(B)\, P(E \mid B) + P(C)\, P(E \mid C). \qquad \text{[Multi-Stage } p\text{-Rule]}$$

$$= \frac{4}{9} \cdot \frac{3}{10} + \frac{2}{9} \cdot \frac{5}{10} + \frac{3}{9} \cdot \frac{8}{10} = \frac{46}{90}.$$

Now by Bayes's Reversal Theorem : $P(A \mid E) = \dfrac{P(A)\, P(E \mid A)}{P(E)} = \dfrac{12}{46}.$

4*. Let the No. of blue and red in the bag be b and r respectively. Let B_k and R_k denote the drawing of k blue and k red balls. Then [draw figure as for hyp-geom model].

$$x = P(R_2) \binom{r}{2}\binom{b}{1} \Big/ \binom{r+b}{2} = \frac{r(r-1)}{(r+b)(r+b-1)}.$$

$$y = P(B_2) \binom{b}{2}\binom{r}{0} \Big/ \binom{r+b}{2} = \frac{b(b-1)}{(r+b)(r+b-1)}.$$

$$z = P(B_1 R_1) \binom{r}{1}\binom{b}{1} \Big/ \binom{r+b}{2} = \frac{2rb}{(r+b)(r+b-1)}.$$

By hypotheses $x = 5y$, $z = 6y \Rightarrow r(r-1) = 5b(b-1)$; $2rb = 6b(b-1)$.
Solving : $b = 1, 3$ and then $r = 0, 6$. So the admissible solution is $b = 3$, $r = 6$.

5*. If event A_i denotes the ith urn chosen and 'R' denotes the event of withdrawing red ball, then

$P(A_1) = P(A_2) = P(A_3) = \frac{1}{3}$ [Urn-selection is equiprobable]

$P(R \mid A_1) = \frac{3}{8}$, $P(R \mid A_2) = \frac{1}{6}$, $P(R \mid A_3) = 4/9$.

$\therefore$ $P(R) = P(A_1)p(R \mid A_1) + P(A_2)P(R \mid A_2) + P(A_3)P(R \mid A_3) = 71/216$. [Multi-stage p-Rule]

So $P(A_2 \mid R) = P(R \mid A_2)\ P(A_2)/P(R) = (1/18)\ (216/71) = 12/71$. [Baye's Reversal Rule]

6*. Denote by T, E, M, F the events that student wears a tie, comes from East, comes from the Mid-West and comes from the Far-West, respectively. The given data now reads .

$P(E) = 0.50$, $P(M) = 0.30$, $P(F) = 0.20$; $P(T \mid E) = 0.80$, $P(T \mid M) = 0.60$, $P(T \mid F) = 0.40$

$\therefore$ $P(T) = P(E)\ P(T \mid E) + P(M)\ P(T \mid M) + P(F)\ P(T \mid F)$ [Multi-stage p-Rule]

$= 0.5 \times 0.8 + 0.3 \times 0.6 + 0.2 \times 0.4 = 0.66$.

We now utilize Baye's Reversal Rule to get

$P(E \mid T) = P(E)\ P(T \mid E)/P(T) = 0.8 \times 0.5/(0.66) = 20/33$.

$P(M \mid T) = P(M)\ P(T \mid M)/P(T) = 0.6 \times 0.3/(0.66) = 3/11$.

$P(F \mid T) = P(F)\ P(T \mid F)/P(T) = 0.4 \times 0.2/(0.66) = 4/33$.

7*. Let C_i $(i = 1, 2, 3)$ denote the event of choosing the ith chest and R the event of getting red ball. Then, $P(C_i) = \frac{1}{3}$. $P\{R \mid C_1\} = 0$, $P\{R \mid C_2\} = \frac{1}{2}$, $P\{R \mid C_3\} = 1$.

The ball obtained being red, the other ball can be red only if the chest is 3rd (C_3) and thus, we require $P\{C_3 \mid R\}$. Now

$P\{R\} = P(C_1)\ P(R \mid C_1) + P(C_2)\ P(R \mid C_2) + P(C_3)\ P(R \mid C_3) = \frac{1}{2}$. [Multi-Stage p-Rule]

$\therefore$ $P(C_3 \mid R) = P(R \mid C_3)\ P(C_3)/\ P\{R\} = \frac{1}{3}.1/\frac{1}{2} = \frac{2}{3}$ [Baye's Reversal Rule]

Remark. To find the probability that the other ball in the chest is black (B), we require $P\{B_2 \mid R\}$.

$$P\{B_2 \mid R\} = P(R \mid B_2)\ P(B_2)/P\{R\} = \tfrac{1}{6}/\left(\tfrac{1}{2}\right) = \tfrac{1}{3}.$$

[You may call balls as gold coins, silver coins, etc. as some Authors do.]

8*. Let S = Statement made ; T = Statement S is true, F = Statement S is false,

C = A and B make same statement S ; D = A asserts S and B denies it.

Then $P(T) = P(F) = \frac{1}{2}$, $P(C \mid T) = pp'$, $P(C \mid F) = qq'$, $P(D \mid T) = qp'$, $P(D \mid F) = p'q$

where$p + q = 1$, $p' + q' = 1$. Now

$P(C) = P(T)\ P(C \mid T) + P(F)\ P(C \mid F) = \frac{1}{2}\ (pp' + qq')$. [Multi-stage p-Rule]

$\therefore P(T \mid C) = P(T)\ P(C \mid T)/P(C) = \frac{1}{2}pp' / \frac{1}{2}(pp' + qq') = pp' / (pp' + qq')$. [Baye's Reversal Rule]

Further, $P(D) = P(T)\ P(D \mid T) + P(F)\ P(D \mid F) = \frac{1}{2}(pq' + p'p)$. [Multi-stage p-Rule]

If A_t denotes the event that A tells the truth, then by Baye's Reversal Rule

$P(A_t \mid D) = P(A_t)\ P(D \mid A_t)/P(D) = qp'/(pq' + qp')$.

9*. Let A, B, C denote the events of selecting these very boxes, W_1 the event of drawing a white chip and W_2 the future event of drawing another white chip from the urn. We are then having :

$P(A) = P(B) = P(C) = \frac{1}{3}$, $P(W_1 \mid A) = \frac{1}{3}$, $P(W_1 \mid B) = \frac{2}{3}$, $P(W_1 \mid C) = \frac{2}{4}$.

$\therefore P(W_1) = P(A)\,P(W_1 \mid A) + P(B)\,P(W_1 \mid B) + P(C)\,P(W_1 \mid C) = \frac{1}{3}\left(\frac{1}{3} + \frac{2}{3} + \frac{2}{4}\right) = \frac{1}{2}$. [Multi-stage p-Rule]

Further, $P(W_2 \mid AW_1) = 0$, $P(W_2 \mid BW_1) = \frac{1}{2}$, $P(W_2 \mid CW_1) = 1/3$.

[$W_2 \mid BW_1$: means to draw white ball from B again]

$$\therefore P(W_2 \mid W_1) = \frac{P(W_2\, W_1)}{P(W_1)} = \frac{\Sigma P(A)\, P(W_1 \mid A)\,(W_2 \mid AW_1)}{\Sigma P(A)\, P(W_1 \mid A)} = \frac{\frac{1}{3}\left(\frac{1}{3}.0 + \frac{2}{3}.\frac{1}{2} + \frac{2}{4}.\frac{1}{3}\right)}{\frac{1}{2}} = \frac{1}{3}. \text{ [Baye's Rule]}$$

10*. Let R denote the event of choosing red ball. Then $P(A) = 2/6$, $P(B) = 4/6$, $P\{R \mid \text{six}\} = P\{R \mid A\} = 4/12$.

$$P\{R\} = P(A)\ P(R \mid A) + P(B)\ P(R \mid B) = \frac{2}{6}.\frac{4}{12} + \frac{4}{6}.\frac{3}{8} = \frac{13}{36}. \qquad \text{[Multi-stage } p\text{-Rule]}$$

$$\therefore\ P\,\{6 \mid R\} = P\{6\}\ P\{R \mid 6\}/P\{R\} = \frac{1}{6}.\frac{4}{12}\left(\frac{13}{36}\right) = \frac{2}{15}. \qquad \text{[Baye's Reversal Rule]}$$

11*. Let T_1 and T_2 denote the events that the transferred ball is white and transferred ball is black respectively. Let W denote the event that a white ball is extracted from B. Then,

$P(T_1) = \frac{2}{4}$, $P(T_2) = \frac{2}{4}$, $P(W \mid T_1) = \frac{4}{6}$, $P(W \mid T_2) = \frac{3}{6}$.

$\therefore\ P(W) = P(T_1)\ (W \mid T_1) + P(T_2)\ P(W \mid T_2) = \frac{7}{12}$. [Multi-stage p-Rule]

$P\{T_1 \mid W\} = P\{W \mid T_1\}\ P\{T_1\}/P\{W\} = \left(\frac{1}{3}\right)/\left(\frac{7}{12}\right) = \frac{4}{7}$. [Baye's Reversal Rule]

12*. Let M, F, S denote the events that all n people chosen are male, female, and are of same the sex, respectively. Then,

$$P(M) = \binom{3n-1}{n}\Big/\binom{6n-1}{n},\ P(F) = \binom{3n}{n}\Big/\binom{6n-1}{n};\ P(S \mid M) = 1 = P(S \mid F).$$

$P(S) = P(M)\ P(S \mid M) + P(F)\ P(S \mid F)$ [Multi-stage p-Rule]

$$= \lambda\binom{3n-1}{n}1 + \lambda\binom{3n}{n}.1. \quad \lambda = \left[\binom{6n-1}{n}\right]^{-1}$$

$\therefore\ P(F \mid S) = P(S \mid F)\ .\ P(F)/P(S)$ [by Baye's Reversal Rule]

$$= \lambda\binom{3n}{n}\Big/\lambda\left[\binom{3n-1}{n} + \binom{3n}{n}\right] = \frac{3}{5}. \qquad \text{(Independent of } n\text{)}$$

13*. Let W_i denote the event that i white balls and hence $(4 - i)$ black balls are transferred, $i = 0, 1, 2, 3$. Thus $P(W_i) = \binom{3}{i}\binom{5}{4-i}\Big/\binom{8}{4}$.

Let W be the event that a white ball is taken out of second (originally empty) vessel. Then

$$P(W_0) = \binom{5}{4}\Big/\binom{8}{4} = \frac{5}{70}, \qquad P(W_1) = \binom{5}{3}\binom{3}{1}\Big/\binom{8}{4} = \frac{30}{70},$$

$P(W_2) = \binom{3}{2}\binom{5}{2}\Big/\binom{8}{4} = \frac{30}{70}$, $\quad P(W_3) = \binom{5}{1}\binom{3}{3}\Big/\binom{8}{4} = \frac{5}{70}$, $\quad P(W_4) = 0$.

$P(W \mid W_0) = 0$, $P(W \mid W_1) = \frac{1}{4}$, $P(W \mid W_2) = \frac{2}{4}$, $P(W \mid W_3) = \frac{3}{4}$, $\quad P(W \mid W_4) = 0$.

$\therefore \quad P(W) = \Sigma P(W_i)\, P(W \mid W_i)$, $i = 0, 1, 2, 3, 4$ [Multi-stage p-Rule]

$$= \frac{5}{70}.0 + \frac{30}{70}.\frac{1}{4} + \frac{30}{70}.\frac{2}{4} + \frac{5}{70}.\frac{3}{4} + 0 = \frac{105}{280} = \frac{3}{8}.$$

So, $\quad P\{W_3 \mid W\} = P(W_3)\, P(W \mid W_3)/P(W) = \frac{5}{70}.\frac{3}{4}\Big/\frac{3}{8} = \frac{1}{7}$. [Baye's Reversal Rule]

14*. Let B_1 and C_k denote the respective events that a family has one boy and a family has k children. We are given : $P(C_k) = P_k$, $[P(C_1 = P_1]$, etc. and we require $P\{C_1 \mid B_1\}$. Now,

$P\{B_1 \mid C_k\}$ = Prob. {a family with k children has just one boy} $= \binom{k}{1} q^{k-1} p$.

$P\{B_1\} = P(C_1)\, P(B_1 \mid C_1) + P(C_2)\, P(B_1 \mid C_2) + \ldots + P(C_n)\, P(B_1 \mid C_n)$ [Multi-stage p-Rule]

$$= \sum_{k=1}^{n} P_k \cdot \binom{k}{1} q^{k-1}\, p = p \sum_{k=1}^{n} k\, P_k\, q^{k-1}. \qquad [P(B_1 \mid C_1) = p]$$

So, $P\{C_1 \mid B_1\} = P\{B_1 \mid C_1\}\, P(C_1)/P(B_1) = P_1/(\Sigma k P_k\, q_{k-1}$. [Baye's Reversal Rule]

15*. Let R be the event that a red ball is selected from box B. Let R_2, B_2, M denote that transferred balls are both red, both blue or one-red and one-blue (mixed). Then

$P\{R\} = P(R_2)\, P(R \mid R_2) + P(B_2)\, P(R \mid B_2) + P(M)\, P(R \mid M)$. [Multi-stage p-Rule] ...(1)

Here $P(R_2) = \binom{4}{2}\Big/N$, $P(B_2) = \binom{8}{2}\Big/N$, $P(M) = \binom{4}{1}\binom{8}{1}\Big/N$. $N = \binom{12}{2} = 66$.

$\therefore \quad P(R \mid R_2) = 8/13$, $P(R \mid B_2) = 6/13$, $P(R \mid M) = 7/13$

$P\{R\} = (6 \times 8 + 28 \times 6 + 32 \times 7)/66 \times 13 = 220/429$. [By (1)]

(ii) L_1 = {atleast one red ball was transferred} $= R_1 \uplus R_2$

$\therefore \quad P\{L_1 \mid R\} = P\{(R_1 \uplus R_2) \mid R\} = P(R_1 \mid R) + P(R_2 \mid R)$ [Use Baye's Resersal Rule]

$$= \frac{P(R_1)\, P(R \mid R_1)}{P(R)} + \frac{P(R_2)\, P(R \mid R_2)}{P(R)} = \frac{(4/12)(7/13)}{220/429} + \frac{(6/66)(8/13)}{220/429} = \frac{77}{220} + \frac{6}{55} = \frac{101}{220}.$$

16*. Define the events : E_A = {even-numbered card from urn A},

OB = {odd-numbered card from B}, E_i = {ith card even}, O_i = [ith card odd}

$\therefore \quad P(E_A) = \frac{4}{9}$, $P(E_B) = \frac{2}{5}$, $\quad P(O_A) = \frac{5}{9}$, $P(O_B) = \frac{3}{5}$.

(i) $P(E_1\, E_2) = P(A)\, P(E_1 E_2 \mid A)) + P(B)\, P(E_1\, E_2 \mid B)$ [Multi-stage p-Rule]

$= \frac{1}{2}.\frac{4}{9}.\frac{3}{8} + \frac{1}{2}.\frac{2}{5}.\frac{1}{4} = \frac{2}{15}$ $\quad [\because P(E_1\, E_2) = P(E_1)\, P(E_2 \mid E_1)]$

(ii) $P(O_1 O_2) = P(A)\, P(O_1 O_2 \mid A) + P(B)\, P(O_1 O_2 \mid B)$ [Multi-stage p-Rule]

$= \frac{1}{2}.\frac{5}{9}.\frac{3}{5} + \frac{1}{2}.\frac{3}{5}.\frac{5}{9} = \frac{1}{3}$ $\quad [\because P(O_1 O_2) = P(O_1)\, P(O_2 \mid O_1)]$

$\therefore \quad P\{A \mid E_1E_2\} = P(A) \,.\, P(E_1E_2 \mid A)\, P(E_1E_2) = \left(\frac{1}{12}\right)\left(\frac{15}{2}\right)\frac{5}{8}$ [Baye's Reversal Law]

17*. Let W_i denote the event that i white (hence $4-i$ black) balls are transferred to second urn and W denote the event that ball from second urn is white (on the first draw). Now

$$P(W_i) = \binom{5}{i}\binom{5}{4-i}\Big/\binom{10}{4} = \lambda\binom{5}{i}\binom{5}{4-i} \text{ where } \lambda^{-1}\ \binom{10}{4} = 210.$$

$$\therefore \quad P(W) = 5\lambda,\ P(W_1) = 50\lambda,\ P(W_2) = 100\lambda,\ P(W_3) = 50\lambda,\ P(W_4) = 5\lambda$$

$$P\{W \mid W_0\} = 0,\ P\{W \mid W_1\} = \tfrac{1}{4},\ P\{W \mid W_2\} = \tfrac{2}{4},\ P\{W \mid W_3\} = \tfrac{3}{4},\ P\{W \mid W_4\} = \tfrac{4}{4}.$$

$$\therefore \quad P(W) = \Sigma i\ P(W_i)\ P(W \mid W_i), \qquad 0 \le i \le 4 \qquad \text{[Multi-stage } p\text{-Rule]}$$

$$= 0 + \tfrac{5}{21}\cdot\tfrac{1}{4} + \tfrac{10}{21}\cdot\tfrac{1}{2} + \tfrac{5}{21}\cdot\tfrac{3}{4} + \tfrac{1}{42}\cdot 1 = \tfrac{1}{2}$$

Let W' denote the event that a second white ball is drawn from the second urn on the second trial. Then,

$$P\{W' \mid WW_0) = 0,\ P(W' \mid WW_1) = 0,\ P(W' \mid WW_2) = \tfrac{1}{3},\ \ P(W' \mid WW_3) = \tfrac{2}{3},\ \ P(W' \mid WW_4) = \tfrac{3}{3}.$$

Now Baya's Reversal law for the **future** event W' is

$$P(W' \mid W) = \{\Sigma_i\, P(W_i)\, P(W \mid W_i)\, P(W' \mid WW_i)]/P(W)$$

$$\text{Num.} = \frac{5}{210}.0.0 + \frac{5}{21}\cdot\frac{1}{4}.0 + \frac{10}{21}\cdot\frac{1}{2}\cdot\frac{1}{3} + \frac{5}{21}\cdot\frac{3}{4}\frac{2}{3} + \frac{1}{42}.1.1 = \frac{2}{9}.$$

Since $\quad$ Den. $= P(W) = \frac{1}{2}$, hence $P(W' \mid W) = \frac{4}{9}$.

Sec. 2.60. Page 95

1*. Here, $n(\Omega) = 36$ are equiprobable possible pairs of outcomes each with prob. 1/36. Now,

$$n(A) = \{(1, 3, 5) \times (1, 2, 3, 4, 5, 6)] = 18 = [(1, 2, 3, 4, 5, 6) \times (1, 3, 5)] = n(B).$$

To find $n(C)$, we note that the sum of eyes is odd, if one die shows odd number and the other die shows even number ; so that

$$n(C) = [(1, 3, 5) \times (2, 4, 6) + (2, 4, 6) \times (1, 3, 5)] = 18.$$

Also, cases favourable to events $A \cap B$, $A \cap C$, $B \cap C$ are

$$n(A \cap B) = [(1, 3, 5) \times (1, 3, 5)] = 9\ ;\ n(B \cap C) = [(2, 4, 6) \times (1, 3, 5)] = 9\ ;$$

$$n(C \cap A) = [(1\ 3, 5) \times (2, 4, 6)] = 9.$$

$A \cap B \cap C = \varnothing$ (Since if two faces read odd, their sum cannot be odd). Hence $n(ABC) = 0$. Thus,

$$P(A) = P(B) = P(C) = 18/36 = 1/2\ ;\ P(AB) = P(BC) = P(CA) = 9/36 = 1/4\ ;\ P(ABC) = 0.$$

$$\therefore\ P(AB) = P(A)\,P(B),\ P(BC) = P(B)\,P(C)\ ;\ P(CA) = P(C)\,P(A)\ ;\ P(ABC) \ne P(A)\,P(B)\,P(C).$$

It follows that the events are *pair-wise independent* but not *mutually independent*. Observe that

$AB = \{(1, 1), (1, 3), (1, 5), (3, 1), (3, 3), (3, 5), (5, 1), (5, 3), (5, 5)\}$

$BC = \{(2, 1), (4, 1), (6, 1), (2, 3), (4, 3), (6, 3), (2, 5), (4, 5), (6, 5)\}$

$CA = \{(1, 2), (1, 4), (1, 6), (3, 2), (3, 4), (3, 6), (5, 2), (5, 4), (5, 6)\}$.

2*. Let D denote the event that 2-headed (dishonest) coin is selected and H denote the event that a head turns up when the coin is tossed. Then $P(D) = P(\bar{D}) = \frac{1}{2}$. Note that, if coin is 2-headed, H is *certain* to occur, but if the coin is fair, H occurs with probability 1/2. Hence, $P(H \mid D) = 1$ and $P(H \mid \bar{D}) = \frac{1}{2}$. Now,

$$P(H) = P(D)\,P(H \mid D) + P(\bar{D})\,P(H \mid \bar{D}) = \tfrac{1}{2}.1 + \tfrac{1}{2}.\tfrac{1}{2} = \tfrac{3}{4}. \qquad \text{[Multi-stage } p\text{-Rule]}$$

$$\therefore \quad P(D \mid H) = \frac{P(D).P(H \mid D)}{P(H)} = \frac{\left(\frac{1}{2}\right)(1)}{\frac{3}{4}} = \frac{2}{3}. \qquad \text{[Baye's Reversal Rule]}$$

3*. Chances of the animal being shot dead at the successive ranges are $(1/2)^2$, $(1/3)^2$, $(1/4)^2$, ..., $(1/n)^2$. The corresponding chances of the animal's escape are

$1 - (1/2)^2,\ 1 - (1/3)^2,\ 1 - (1/4)^2,\ \ldots,\ 1 - (1/n^2)$.

Thus, the chance of escape of the animal at the end ($r = na$) (by Product Rule) is,

$$p = \left(1 - \frac{1}{2^2}\right)\left(1 - \frac{1}{3^2}\right)\left(1 - \frac{1}{4^2}\right)\ldots\left(1 - \frac{1}{n^2}\right) = \prod_{x=2}^{n}\left(1 - \frac{1}{x^2}\right)$$

$$= \prod_{x=2}^{n}\left(\frac{x-1}{x}\right).\left(\frac{x+1}{x}\right) = \left(\frac{1}{2}.\frac{3}{2}\right)\left(\frac{2}{4}.\frac{4}{3}\right)\left(\frac{3}{5}.\frac{5}{4}\right)\ldots\left(\frac{n-1}{n}.\frac{n+1}{n}\right) = \frac{n+1}{2n}.$$

So
$$q = 1 - p = 1 - \left(\frac{n+1}{2n}\right) = \frac{n-1}{2n}. \qquad \text{(Chance of animal being killed)}$$

Thus the Odds against the Sportsman are $p : q = n + 1 : n - 1$.

4*. Let $p = \frac{1}{3}$. Obviously, B wins in three more games with probability $\left(\frac{1}{3}\right)^3 = p^3$.
$P(B$ wins 4th game$) = [3p^3 + 3p^3]/3$, $P(B$ wins 5th game$) = [6p^4 + 12p^4]/3$,
$P(B$ wins 6th game$) = 30p^5/3$. Thus $P(B$ wins$) = 55/243$ (Total probability).

Chapter 3 : One-Variate Distribution Theory

Sec. 3-21. Page 119

1*. We know that a c.d.f. has to be non-negative. Now if $0 < x < 1$, then $\ln x < 0$, this necessitates that k be negative. Also if $1 < x < 3$, then $\ln x > 0$ and this does require that k be positive. It means there is no constant k which would make $k \ln x \geq 0$ for all x in the interval $]\,0, 3\,[$. Thus, F cannot represent a c.d.f. for any non-zero constant k. If $k = 0$, then F is a c.d.f.

2*. (i) If $X < a$, then $\{X \leq a\} = \varnothing$, $\Rightarrow P(X \leq x) = P(\varnothing) \Rightarrow F(x) = 0$.

If $X > b$, then $\{X \leq a\} = \Omega$, $\Rightarrow P(X \leq x) = P(\Omega) \Rightarrow F(x) = 1$.

(ii) If $Y(\omega) \leq t$, then $X(\omega) \leq t$, since $X(\omega) \leq Y(\omega)$.

$\therefore \quad \{Y \leq t\} \subset \{X \leq t\} \Rightarrow P(Y \leq t) \leq P(X \leq t) \Rightarrow F_Y(t) \leq F_X(t)$.

Sec. 3-32. Page 122

1*. Here X is discrete and $P\{X = x\} = k(x + 1)(1/5)^x$. Since $\Sigma p(x) = 1$, $x = 0, 1, 2, \dots$ we get, putting $1/5 = r$.

$$1 = k\Sigma (x + 1) r^x = k(1 + 2r + 3r^2 + 4r^3 + \dots) = k(1 - r)^{-2} = k\,(25/16) \Rightarrow k = 16/25$$

$$\therefore \quad p(x) = (16/25)(x + 1)(1/5)^x, \; x = 0, 1, 2, \dots$$

$$P(X \le 5) = P(X = 0) + P(X = 1) + \dots + P(X = 5) = \frac{16}{25}\left[1 + \frac{2}{5} + \frac{3}{5^2} + \dots \frac{6}{5^5}\right]$$

$$= 0.64\,(1 + 0.4 + 0.12 + 0.032 + 0.008 + 0.00192) = 0.64 \times 1.562 = 0.9997$$

Note. Since $F(5) = 0.9997$, it follows that, for $X > 5$, $F(x)$ will change very slowly

2*. We find c.d.f. of X. Thus, since $F(x) = P(X \le x)$

$$\therefore \quad F(x) = \frac{2}{(n+1)(n+1)} \sum_{y=0}^{x} (y+1) = \frac{(x+1)(x+2)}{(n+1)(n+2)} \qquad \dots(1)$$

Thus, $\quad P(A) = P(X \le m) = (m + 1)(m + 2)/(n + 1)(n + 2). \qquad \dots(2)$

An appropriate model conditioned on event A for $x \in A$, is

$$P\{X \mid A\} = P\{X, A\}/P(A) = p(x)/P(A) = 2\,(x + 1)/(m + 1)(m + 2), \; x = 0, 1, 2, \dots, m.$$

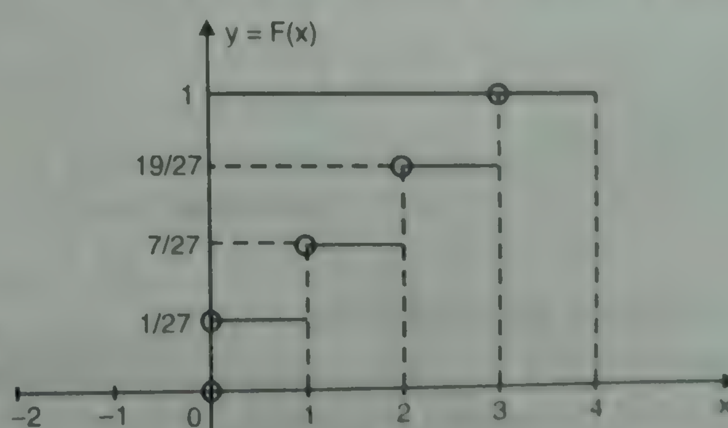
$$P\{B \mid A\} = P(B \cap A)/P(A) = P(B)/P(A) = F(l)/F(m) = (l + 1)(l + 2)/(m + 1)(m + 2). \qquad [\because B \subset A]$$

3*. The sample space, random values and probabilities are as follows :

Ω :	$\{(hhh)$,	(hht),	(hth),	(thh),	(htt),	(tht),	(tth),	$(ttt)\}$
X :	3	2	2	2	1	1	1	0
p :	$\frac{8}{27}$	$\frac{4}{27}$	$\frac{4}{27}$	$\frac{4}{27}$	$\frac{2}{27}$	$\frac{2}{27}$	$\frac{2}{27}$	$\frac{1}{27}$

$F_X(x) = P(X \le x\}$, so we get

$$P(X \le x) = \begin{cases} 0, & x < 0 \\ 1/27, & 0 \le x < 1 \\ 7/27, & 1 \le x < 2 \\ 19/27, & 2 \le x < 3 \\ 1, & 3 \le x \end{cases}$$



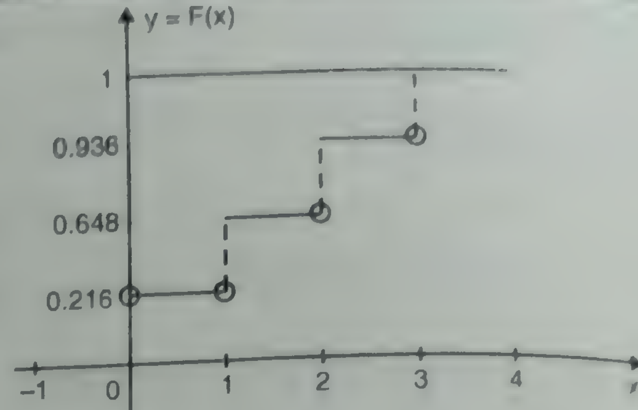
The circle on the graph are used to indicate the right-continuity of the function $F(x)$.

4*. Let X denote the number of senior teachers amongst the three members chosen, then $\Omega = \{(x_1, x_2, x_3) : x_i = 1 \text{ or } 0\}$. Here X_i are indicators and $X = X_1 + X_2 + X_3$ and range of X is $\{0, 1, 2, 3\}$, so that X is a discrete variate. Now

$$P\{(0, 0, 0)\} = (0.6)^3 = 0.216 \; ; \; P\{(1, 0, 0)\}$$

$$= (0.4)(0.6)^2 = 0.144, \text{ etc.}$$

$$\therefore \quad p(0) = P(X = 0) = P(0, 0, 0) = 0.216$$



$p(1) = P(X = 1) = P\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$

$= 3 \times 0.144 = 0.432$

$p(2) = P(X = 2) = P\{(1, 1, 0), (1, 0, 1), (0, 1, 1)\}$

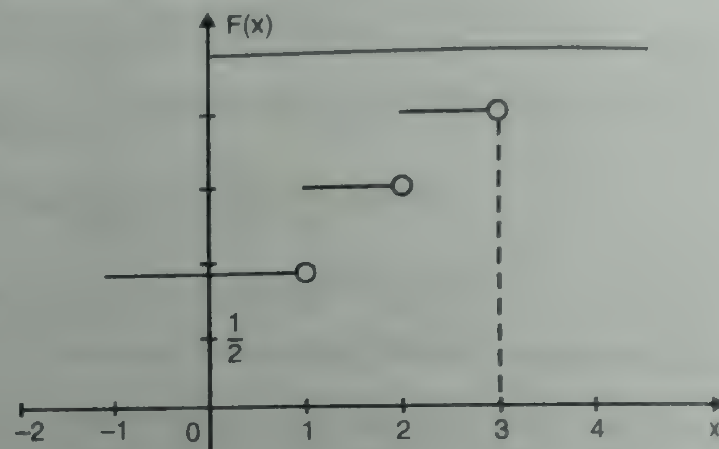
$= 3 \times (0.4)^2\,(0.6) = 0.288$

$p(3) = P(X = 3) = P\{(1, 1, 1)\} = (0.4)^3 = 0.064$

The distribution function $F_X(x) = P(X \leq x)$ is

$F_X(x) = 0\ I\ (x < 0) + 0.216\ I(x \leq 0) + 0.648\ I(x \leq 1) + 0.936\ I(x \leq 2) + 1\ I\ (x \leq 3)$.

5*. $\Omega = \{(x_1, x_2, x_3)\ ;\ x_j = 1, 2, \ldots, 6, j = 1, 2, 3\}$.



$P\{\text{integer is } k\} = \frac{1}{6} = p$;

$P\{\text{integer is not } k\} = \frac{5}{6} = q$.

Range of $X = \{-1, 1, 2, 3\}$

$P(X = -1) = q^3 = 125/216$;

$P(X = 1) = 3pq^2 = 75/216$

$P(X = 2) = 3p^2q = 15/216$;

$P(X = 3) = p^3 = 1/216$.

Note that $X = 2$ means $(k, k, 0) \uplus (k, 0, k) \uplus (0, k, k)$, etc. Thus,

$$f_X(x) = \begin{cases} 125/216 & \text{for } x = -1 \\ 75/216 & \text{for } x = 1 \\ 15/216 & \text{for } x = 2 \\ 1/216 & \text{for } x = 3 \end{cases}.$$

The distribution function is $F(x) = P(X \leq x)$. Now

$$F_X(x) = \begin{cases} 0 & x < -1 \\ 125/216 & x \leq -1 \\ 200/216 & x \leq 1 \\ 215/216 & x \leq 2 \\ 1 & x \leq 3 \end{cases}.$$

Evaluation of probabilities :

(a) $P(0 < X \leq 3) = F(3) - F(0) = 1 - (125/216) = 91/216)$.

(b) $P(X \leq 0) = F(0) = 125/216$.

(c) $P(-1 < X \leq 0) = F(0) - F(-1) = (125/216) - (125/216) = 0$.

6*. We require to evaluate $\Sigma\ (e^{a+b} - 1)^{-1}$.

Let $A_1 = \{a, b : (a, b) = 1\}$. Then, putting $m = a + b$

$$\Sigma_{A_1}(e^{a+b} - 1)^{-1} = \Sigma_{A_1} e^{-m}(1 - e^{-m})^{-1} = \Sigma_{A_1} e^{-m}(1 + e^{-m} + e^{-2m} + e^{-3m} + \ldots)$$

$$\Sigma_{A_1} e^{-m} + \Sigma_{A_1} e^{-2m} + \Sigma_{A_1} e^{-3m} + \ldots = \Sigma_{A_1} e^{-m} + \Sigma_{A_2} e^{-2m} + \Sigma_{A_3} e^{-3m} + \ldots$$

where $A_n = \{a, b : (a, b) = n\}$. Thus, the sum can be replaced by $\Sigma_A e^{-n}$, where A is given by $A = A_1 \cup A_2 \cup \ldots$ and as such A is the set of all positive integers. Thus,

$$\sum_{A_1}(e^{a+b}-1)^{-1} = \sum_{a,b} e^{-(a+b)} = \left(\sum_{a=1}^{\infty} e^{-a}\right)\left(\sum_{b=1}^{\infty} e^{-b}\right) = \left(\frac{e^{-1}}{1-e^{-1}}\right)^2 = (e-1)^{-2}$$

Observe that the series involved are absolutely convergent so that the double sum has been expressed as a repeated sum. Thus, $K\,\Sigma(e^{a+b}-1)^{-1} = 1 \Rightarrow K = (e-1)^2$

Now $X = 1 \Leftrightarrow a = b, (a, b) = 1 \Leftrightarrow a = b = 1$. Hence,

$$P(X = 1) = K(e^2-1)^{-1} = (e-1)^2/(e^2-1) = (e-1)/(e+1).$$

$$P(X < 1) = \Sigma_B\, P(X = a/b),\ B = \{a, b : (a, b) = 1, a < b\}.$$

$$P(X > 1) = \Sigma_C\, P(X = a/b),\ C = \{a, b : (a, b) = 1, a > b\}.$$

Since $B = C$, it follows that $P(X < 1) = P(X > 1) = p$ (say). Now,

$P(X < 1) + P(X = 1) + P(X > 1) = 1 \Rightarrow 2p = 1 - (e-1)/(e+1) = 2/(e+1) \Rightarrow p = P(X < 1) = 1/(e+1)$

$P(X \le 1) = P(X < 1) + P(X = 1) = [1/(e+1)] + [(e-1)/(e+1)] = e/(e+1).$

Sec. 3-55. Page 131

1*. The distribution function F is not differentiable at the sharp edges : $x = 0$, $x = 1$ Except for these points, F is differentiable everywhere. Differentiating F we get

$$f(x) = 0,\ x < 0, \text{ or } x > 1\ ;\ f(x) = 1,\ 0 < x < 1.$$

At $x = 0, 1$ we are free to choose $f(0)$, $f(1)$ arbitrarily. Let us choose $f(0) = f(1) = 1$, and thus we obtain

$$f(x) = 1,\ 0 \le x \le 1\ ;\ f(x) = 0, \text{ elsewhere.}$$

$P(\frac{1}{3} \le X \le \frac{1}{2}) = F(\frac{1}{2}) - F(\frac{1}{3}) = \frac{1}{2} - \frac{1}{3} = \frac{1}{6}$. $\qquad P(\frac{1}{2} \le X \le 2) = F(2) - F(\frac{1}{2}) = 1 - \frac{1}{2} = \frac{1}{2}$.

Using density function f we get

$$P\left(\tfrac{1}{3} \le X \le \tfrac{1}{2}\right) = \int_{1/3}^{1/2} f(x)\,dx = \int_{1/3}^{1/2} dx = \tfrac{1}{2} - \tfrac{1}{3} = \tfrac{1}{6}.$$

$$P\left(\tfrac{1}{2} \le X \le 2\right) = \int_{1/2}^{2} f(x)\,dx = \int_{1/2}^{1} 1\,dx = \tfrac{1}{2}, \qquad [f(x) = 0, \text{ when } x > 1].$$

2*. The c.d.f. 'F' is not differentiable at $x = 0$, $x = \frac{1}{2}$, $x = 3$ (sharp edges) and except for these points, f is differentiable everywhere. Differentiating F we get

$$f(x) = \begin{cases} 0, & x < 0 \text{ or } x > 3 \\ 2x, & 0 < x < \frac{1}{2} \\ 6(3-x)/25 & \frac{1}{2} < x < 3 \end{cases}$$

At $0, \frac{1}{2}, 3$ we are free to choose $f(0), f(\frac{1}{2}), f(3)$ arbitrarily. Let us choose $f(0) = 0$, $f(\frac{1}{2}) = \frac{3}{5}$, $f(3) = 0$; then

$f(x) = 2x,\ 0 \le x < \frac{1}{2}\ ;\ f(x) = 6(3-x)/25,\ \frac{1}{2} \le x < 3\ ;\ f(x) = 0$, elsewhere.

(i) $P(|X| \le 1) = P(-1 \le X \le 1) = \displaystyle\int_0^{1/2} 2x\,dx + \int_{1/2}^{1} \frac{6}{25}(3-x)\,dx = \frac{1}{4} + \frac{27}{100} = \frac{52}{100} = \frac{13}{25}$

where we have used $f(x) = 0$ for $x < 0$; and have split the range.

(ii) $P\left(\frac{1}{3} \le X < 4\right) = \int_{1/3}^{3} f(x)\,dx = \int_{1/3}^{1/2} (2x)\,dx + \int_{1/2}^{3} \frac{6}{25}(3-x)\,dx = \frac{8}{9}$

(i) $P(-1 \le X \le 1) = F(1) - F(-1) = \left[1 - \frac{3}{25}(3-1)^2\right] - 0 = \frac{13}{25}$.

(ii) $P(\frac{1}{3} \le X < 4) = F(4) - F(\frac{1}{3}) = 1 - (\frac{1}{9}) = \frac{8}{9}$.

3*. Obviously F is discontinuous at $x = 0, 2, 3$. Hence

$$p(x) = \begin{cases} F(0) - F(0-0) = \frac{1}{2}, & \text{at } x = 0 \\ F(2) - F(2-0) = \frac{5}{6} - \frac{1}{2} = \frac{1}{3}, & \text{at } x = 2 \\ F(3) - F(3-0) = 1 - \frac{5}{6} = \frac{1}{6}, & \text{at } x = 3 \end{cases}$$

That is, $p(0) = \frac{1}{2}, p(2) = \frac{1}{3}, p(3) = \frac{1}{6}$. Here X is discrete and $p(0) + p(2) + p(3) = 1$.

4*. Since X is continuous, F has no jump-discontinuities. As usual
$P(a \le X \le b) = F(b) - F(a) = 0.7 - (0.5) = 0.2$; $P(b \le X \le c) = F(c) - F(b) = 0.8 - 0.7 = 0.1$

$$P(-\infty < X_1 \le a) = F(a) = 0.5 \quad P(c < X_2 < \infty) = 1 - F(c) = 1 - 0.8 = 0.2.$$

Since X_1 and X_2 are independent variates, so $P(X_1 \in A, X_2 \in B) = P(X_1 \in A) \,.\, P(X_2 \in B)$

$\therefore \quad P\{-\infty < X_1 \le a, c < X_2 < \infty\} = P(-\infty < X_1 \le a) \,.\, P(c < X_2 < \infty) = 0.10.$

5*. $\qquad f(x) \ge 0 \Rightarrow kx(1-x) \ge 0 \Rightarrow x(1-x) \ge 0$ since $k > 0$.

This gives $x > 0, x < 1$; $x < 0, x > 1$. The latter result is inadmissible by (1). Hence $0 \le a < x < b \le 1$. To determine the normalization constant k, we have

$$1 = k\int_a^b (x - x)^2\,dx = \tfrac{1}{6}k[3(b^2 - a^2) - 2(b^3 - a^3)].$$

Thus, $\qquad k = 6/(b-a).\ [3(b+a) - 2(a^2 + ab + b^2)]$.

6*. $\qquad F(x) = \int_{-\infty}^{t} f(t)\,dt = \left(\int_{-\infty}^{0} + \int_{0}^{t}\right) f(t)\,dt = \int_{0}^{t} f(t)\,dt$

$\therefore \qquad F(x) = \int_0^x \frac{t}{3}\,dt = \frac{x^2}{6}, 0 < x \le 1$

$$F(x) = \int_0^1 \frac{t}{3}\,dt + \int_1^x \frac{5}{27}.(4-t)\,dt = -\frac{13}{27} + \frac{5}{27}\left(4x - \frac{1}{2}x^2\right),\quad 1 < x < 4.$$

For $x \ge 4$, $F(x) = 1$. Thus, $F(x)$ may be expressed as :

$$F(x) = \begin{cases} 0, & -\infty < x < 0 \\ x^2/6, & 0 < x \le 1 \\ -(13/27) + 5(4x - \frac{1}{2}x^2)/27, & 1 < x \le 4 \\ 1, & x \ge 4 \end{cases}$$

$$F(3) = -(13/27) + (5/27)\,[12 - (9/2)] = 49/54.$$

7*. $\int_{-\infty}^{\infty} f(x)\,dx = k\int_0^{\infty} e^{-kx}\,dx = k\,.\dfrac{\Gamma(1)}{k} = 1.$

where we have used the fact that the integrand is even in the range $]-\infty, \infty[$ and replaced $|x|$ by x when $x \in [0, \infty[$. Trivially, we observe that $f(x) \geq 0 \ \forall\ x \in\ ]-\infty, \infty[$. Thus, f is certainly a density function and X with this density is called "Laplace distribution."

$$F(x) = \int_{-\infty}^{x} f(t)\,dt = \tfrac{1}{2}k\int_{-\infty}^{x} e^{-k|t|}\,dt$$

(i) When $x \leq 0$, $[x = -y,\ y > 0]$ we have

$$F(x) = \tfrac{1}{2}k\int_{-\infty}^{x} e^{-k|t|}\,dt = \tfrac{1}{2}k\int_{-\infty}^{-y} e^{-k|t|}\,dt = \tfrac{1}{2}k\int_{y}^{\infty} e^{-kt}dt = \tfrac{1}{2}e^{-ky} = \tfrac{1}{2}e^{kx}.$$

(ii) When $x > 0$, we have

$$F(x) = \frac{1}{2}k\left(\int_{-\infty}^{0} e^{kt}\,dt + \int_0^{x} e^{-kt}\,dt\right) = 1 - \frac{1}{2}e^{-kx}$$

$$\therefore \qquad F_x(x) = \begin{cases} \tfrac{1}{2}e^{-kx}, & x \leq 0 \\ 1 - \tfrac{1}{2}e^{-kx}, & x > 0 \end{cases}.$$

8*. Here $F(x) = P(X \leq x) = \int_{100}^{x} \dfrac{100}{t^2}\,dt = 1 - \dfrac{100}{x}.$

If p is the probability that a tube lasts for first 150 hours, then

$$p = P(X \leq 150) = F(150) = 1 - (100/150) = 1/3.$$

Hence, the probability that none of three such tubes is to be replaced during the first 150 hours is $p.p.p = p^3$, 1/27 [by independence of tube's lives].

Since $q = 1 - p = \frac{2}{3}$, the probability that all three tubes are burnt out (i.e. have to be replaced) is $q.q.q. = \frac{8}{27}$ [again by independence].

9*. Firstly, we find distribution function of X. Thus

$$F(x) = P(X \leq x) = \theta\int_0^{x}(1+t)^{-\theta-1}\,dt = 1 - \frac{1}{(1+x)^{\theta}}. \qquad ...(1)$$

(a) $\therefore\ P\{X \geq a\} = 1 - F(a) = 1/(1+a)^{\theta}$.

Thus shows that as θ increases, $1/(1+a)^{\theta}$ decreases ; so θ is a sensible m.h. of wood.

(b) $P(X \leq 1) \geq 0.99 \Rightarrow 1 - (\tfrac{1}{2})^{\theta} \geq 0.99$, *i.e.* $2^{-\theta} \leq 0.01$,

$\therefore\ -\theta \ln_{10} 2 \leq -2 \ln 10 \Rightarrow \theta \geq 2/\ln_{10} 2 = 6.6438561.$

Thus, any value of $\theta \geq 6.64$ ensures the requisite.

(c) $p = P\{1 \leq X \leq 2\} = F(2) - F(1) = 2^{-\theta} - 3^{-\theta}$, [by (1)]

$p' = 3^{-\theta}\ln_e 3 - 2^{-1}\ln_e 2\ p'' = 2^{-\theta}(\ln 2)^2 - 3^{-1}(\ln 3)^2$. $[p' = dp/d\theta,\ p'' = d^2p/d\theta^2]$

$p' = 0 \Rightarrow (3/2)^{\theta} = (\ln 3/\ln 2) = L$ say ...(1)

$\therefore \alpha = (\ln L)/\ln (3/2) = \theta_0$, say ...(2)

So $p''(\theta_0) = (3^{-\theta} \ln 3)/(\ln 2/3) < 0$. [by (1)]. Hence p is maximum for θ given by (2).

Sec. 3-73. Page 141

1*. $F_Y(y) = P\{Y \le y\} = P\{\,|X| \le y\} = P(-y \le X \le y\} = P(-y < X \le y) \uplus (X = -y)\}$
$= P\{-y < X \le y\} + P(X = -y) = F_X(y) - F_X(-y) + P(X = -y)$ (Dist. Properties)

Obviously, $F_Y(y) = 0$, if $y \le 0$.

$F_Z(z) = P\{\sqrt{Y} \le z\} = P\{Y \le z^2\} = F_Y(z^2)$. [Obviously, $F_Y(z^2) = 0$, if $z < 0$)]

Now use first part to get

$$F_Z(z) = F_X(z^2) - F_X(-z^2) + P(X = -z^2)\ ;\ F_Z(z) = 0, \quad \text{if } z < 0.$$

Exercise. Find the density of $Y = |X|$ when X is (a) $N(\mu, \sigma^2)$, (b) Lap (λ, a) (c) Chy (a, b).

2*. Let $Y = e^{-X}$, and use transformation $y = e^{-x}$, i.e. $x = -\ln y$.

This gives $|dx/dy| = 1/y$; $x \ge 1 \Rightarrow y \le (1/e)$.

Further $y > 0$ (obviously). From $f_X(x) = |dy/dx|\, f_Y(y)$ we readly get

$f_Y(y) = [y(\ln y)]^{-2}$, $0 < y \le e^{-1}$, $f_Y(y) = 0$, elsewhere.

3*. Let $y = \tan^{-1} x \Rightarrow x = \tan y$ so that $dx/dy = \sec^2 y > 0$.

Thus the function is increasing. Also $1 + x^2 = 1 + \tan^2 y = \sec^2 y$. Hence

$$f_Y(y) = \left[f_x(x) \left| \frac{dx}{dy} \right| \right]_y = \frac{\sec^2 y}{\pi \sec^2 y} = \frac{1}{\pi}, \quad -\frac{\pi}{2} < y < \frac{\pi}{2}.$$

Sec. 3-80. Page 143

1*. The area of the triangle with length a is $\Delta = \frac{1}{2} a^2 \sin(\pi/3) = a^2 \sqrt{\frac{3}{4}}$. Since the area under a probability curve is unity, we get $1 = a^2 \sqrt{\frac{3}{4}}$ or $a = 2(3)^{-1/4}$.

2*. Since $-1 \le \sin(\pi X/2) \le 1$, the only possible values of Y are $-1, 0, 1$. So

$\{Y = 0\} \Leftrightarrow \{X = 2n, n = 1, 2, 3, \ldots\}$, $\{Y = 1\} \Leftrightarrow \{X = 4n + 1, n = 0, 1, 2, \ldots\}$

$\{Y = -1\} \Leftrightarrow \{X = 4n + 3, n = 0, 1, 2, \ldots\}$

$$\therefore\ P(Y = 0) = \sum_{n=1}^{\infty} P(X = 2n) = \sum_{n=1}^{\infty} \left(\frac{1}{2}\right)^{2n} = \sum_{n=1}^{\infty} \left(\frac{1}{2}\right)^{2n} = \frac{\frac{1}{4}}{1 - (\frac{1}{4})} = \frac{1}{3}.$$

$$P(Y = 1) = \sum_{n=0}^{\infty} P(X = 4n + 1) = \sum_{n=0}^{\infty} \left(\frac{1}{2}\right)^{4n+1} = \frac{1}{2} \sum_{n=0}^{\infty} \left(\frac{1}{16}\right)^{n} = \frac{1}{8} \frac{1}{1 - (1/16)} = \frac{8}{15}.$$

$$P(Y = -1) = \sum_{n=0}^{\infty} P(X = 4n + 3) = \sum_{n=0}^{\infty} \left(\frac{1}{2}\right)^{4n+3} = \frac{1}{8} \sum_{n=0}^{\infty} \left(\frac{1}{16}\right)^{n} = \frac{1}{8} \frac{1}{1 - (1/16)} = \frac{2}{15}.$$

3*. Let $y = \sin x$; then $dy/dx = \cos x = \sqrt{1-\sin^2 x} = \sqrt{1-y^2}$. Further

$$x = \pm \pi/2 \Rightarrow y = \pm 1.$$

$$\therefore \quad f_Y(y) = \left[f_X(x) \left| \frac{dx}{dy} \right| \right]_y = \frac{1}{\pi} \frac{1}{\sqrt{1-y^2}}, - < y < 1.$$

4*. Let $y = e^x$, then $x = \ln y$, $(dx/dy) = 1/y > 0$, $[x = 0 \Rightarrow y = 1, x = 1 \Rightarrow y = e]$

$$f_Y(y) = \{f_X(x) \mid dx/dy \mid\}_y = y^{-1}, \quad 1 \le y \le e.$$

Chapter 4 : Jointly Distributed Random Variables

Sec. 4-12. Page 153

1*. Recall $p(x, y) = F(x, y) + F(x^-, y^-) - F(x^-, y) - F(x, y^-)$

where $p(x, y) = P(X = x, Y = y)$. Now

$$p(-2, -5) = F(-2, -5) + F(-2^-, -5^-) - F(-2^-, -5) - F(-2, -5^-) = \tfrac{1}{4} + 0 - 0 - 0 = \tfrac{1}{4}$$

$$p(-2, 3) = F(-2, 3) + F(-2^-, 3^-) - F(-2^-, 3) - F(-2, 3^-) = \tfrac{1}{2} + 0 - \tfrac{1}{4} - 0 = \tfrac{1}{4}$$

$$p(2, -5) = F(2, 5) + F(2^-, -5^-) - F(2^-, -5) - F(2, -5^-) = \tfrac{1}{2} + 0 - 0 - \tfrac{1}{4} = \tfrac{1}{4}$$

$$p(2, 3) = F(2, 3) + F(2^-, 3^-) - F(2^-, 3) - F(2, 3^-) = 1 + \tfrac{1}{4} - \tfrac{1}{2} - \tfrac{1}{2} = \tfrac{1}{4}.$$

2*. Formally : $G(x, \infty) = 1 - e^{-2x}$, $x > 0$; $G(\infty, y) = 0$ or $\frac{1}{2}$ or 1, according as $y < 0$, or $0 < y < 1$ or $y > 1$. These are well-known distribution properties. Now, by the Rectangle Rule :

$$p = P(a < X \le b ; c < Y \le d) = F(b, d) + F(a, c) - F(a, d) - F(b, c).$$

To eliminate exponentials, take $a = \ln 4$, $b = \ln 6$, $c = \frac{1}{4}$, $d = 1$; then

$$p = \left(1 - \frac{1}{36}\right) + \frac{1}{2}\left(1 - \frac{1}{4}\right) - \left(1 - \frac{1}{16}\right) - \frac{1}{2}\left(1 - \frac{1}{6}\right) = \frac{-1}{144} < 0.$$

It follows that $G(x, y)$ cannot be a c.d.f.

Remark. Construction of bivariate c.d.f. is more complicated than its computation.

3*. The required probability p is given by

$$\begin{aligned} p &= P\{1 \le X \le 2, 1 \le Y \le 2\} = F(2, 2) + F(1, 1) - F(1, 2) - F(2, 1) \\ &= 1(1 - a^4 - a^{-8} + a^{-12}) + (1 - a^{-1} - a^{-2} + a^{-3}) - (1 - a^{-1} - a^{-8} + a^{-9}) - (1 - a^{-4} - a^{-2} + a^{-6}) \\ &= a^{-3} - a^{-6} - a^{-9} + a^{-12} = a^{-3}(1 - a^{-3})(1 - a^{-6}). \end{aligned}$$

4*. Let $\quad F(x, y) = 1 - e^{-x} - e^{-y} + e^{-(1/2)M} : x > 0, y > 0, M = \max(x, y) \quad \ldots(1)$

Let $y \to \infty$; then $e^{-y} \to 0$, $e^{-M/2} \to 0$ and so $F_X(x) = 1 - e^{-x}$, $x \ge 0$.

Let $x \to \infty$; then $e^{-x} \to 0$, $e^{-M/2} \to 0$ and so $F_Y(y) = 1 - e^{-y}$, $y \ge 0$.

Now consider : $G(x, y) = 1 - e^{-x} - e^{-y} + (e^x + e^y - 1)^{-1}$; $x, y \ge 0$. $\quad \ldots(2)$

Let $y \to \infty$, then $e^{-y} \to 0$, $(e^x + e^y - 1)^{-1} \to 0$, and so $G_X(x) = 1 - e^{-x}$, $x \geq 0$

Let $x \to \infty$, then $e^{-x} \to 0$, $(e^x + e^y - 1)^{-1} \to 0$, and so $G_Y(y) = 1 - e^{-y}$, $y \geq 0$

Thus, although $F(x, y) \neq G(x, y)$; $G_X(x) = F_X(x)$ and $G_Y(y) = F_Y(y)$.

Note. $F(x, y) = 0$ and $G(x, y) = 0$ for $x, y < 0$. Such values are always understood (implied)

Sec. 4-31. Page 158

1*. Let Y be the number on the second die D_2, then $S = X + Y$ so $Y = S - X$. Now by independent of outcomes on D_1 and D_2.

$P(X = j, S = k) = P(X = j, Y = k - j) = P(X = j)\ P(Y = k - j) = 1/36$, $[1 \leq j \leq 6, 2 \leq k \leq 12]$

Here $1 \leq j < k \leq 12$; $1 \leq k - j \leq 6$. If these inequalities don't hold, $f(j, k) = P(Z) = 0$. Now for outcomes on D_1 $P(X = j) = 1/6$, $1 \leq j \leq 6$.

$$P(S = k) = \sum_{j=1}^{6} P(X = j, S = k).$$

In this sum, the non-zero terms, each 1/36, are those for which either $1 \leq j < k \leq 6$ or $1 \leq k - 6 \leq j \leq 6$. Hence

$$P(S = k) = \begin{cases} (k-1)/36, & \text{if } k \leq 6 \\ (13-k)/36, & \text{if } k > 6. \end{cases} \quad \text{i.e. } P(S_k) = \frac{6 - |7 - k|}{36}, 2 \leq k \leq 2.$$

Remark. Tabulation of outcomes yields $P(S = k)$ visually and convincingly. See Chapter 1

2*. The 2-dice experiment yields the joint p.m.f. which is shown in §1-42. Thus, $(p = 1/36)$

$$f(x, y) = p, \text{ if } 1 \leq x = y \leq 6\ ; \qquad f(x, y) = 2p,\ 1 \leq x < y \leq 6.$$

$$f_2(y) = (2y-1)p,\ 1 \leq y \leq 6 \qquad f_1(x) = (13-2x)p,\ 1 \leq x \leq 6.$$

$$f_1(x \mid y) = f(x, y)/f_2(y)$$

This gives, $$f_1(x \mid y) = \begin{cases} 1/(2y-1), & \text{if } y = x \\ 2/(2y-1), & \text{if } x = 1, 2, \ldots, y-1. \end{cases}$$

Note. For tetrahedral dice, $1 \leq x, y \leq 4$, note that

$$f_1(x) = (9-2x)p, \quad f_2(y \mid x) = \begin{cases} 1/(9-2x), & \text{if } x = y \\ 2/(9-2x), & \text{if } 1 \leq x < y < 6. \end{cases}$$

3*. There are the following numbers of cards in the deck :

Spade-Ace (SA) = 1, Non-Spade Aces (NSA) = 3; Spade-non-Aces (SNA) = 12,

NSNA = 36. Now two cards out of 52 cards can be selected in $\binom{52}{2} = 1326 = k^{-1}$ ways

We layout the plan for calculating the favourable cases as under.

$P(X > Y) = (108k + 3k + 3k) = 114k = 114/1326.$ $[k = 1/1326]$

Values (x, y)	1 SA	3 NSA	12 SNA	36 NSNA	Favourable cases f	Prob. p
(0, 0)	0	0	0	2	$\binom{36}{2} = 630$	630k
(0, 1)	0	0	1	1	$\binom{12}{1}\binom{36}{1} = 432$	432k
(0, 2)	0	0	2	0	$\binom{12}{2} = 66$	66k
(1, 0)	0	1	0	1	$\binom{3}{1}\binom{36}{1} = 108$	108k
(1, 2)	1	0	1	0	$\binom{1}{1}\binom{12}{1} = 12$	12k
(1, 1)	1	0	0	1	$\binom{1}{1}\binom{36}{1} = 36$	72k
	0	1	1	0	$\binom{3}{1}\binom{12}{1} = 36$	
(2, 0)	0	2	0	0	$\binom{3}{2} = 3$	3k
(2, 1)	1	1	0	0	$\binom{1}{1}\binom{3}{1} = 3$	3k
(2, 2)	0	0	0	0	0	0

4*. By normality $\sum p_{ij} = 64a = 1 \Rightarrow 1/64$. The marginal p.m.f's of X and Y are the row and column totals of the given bivariate tabular distribution; thus

$x \downarrow$ \ $y \rightarrow$	1	2	3	4	5	6	Total p_x
0	0	0	$2a$	$4a$	$4a$	$6a$	$16a$
1	$4a$	$4a$	$8a$	$8a$	$8a$	$8a$	$40a$
2	$2a$	$2a$	a	a	0	$2a$	$8a$
Total p_y	$6a$	$6a$	$11a$	$13a$	$12a$	$16a$	$64a$

$\therefore\ P(X = 0) = 16a, \quad P(X = 1) = 40a, P(X = 2) = 8a$;

$P(Y = 1) = 6a, P(Y = 2) = 6a, P(Y = 3) = 11a, P(Y = 4) = 13a, P(Y = 5) = 12a, \ P(Y = 6) = 16a.$

(i) $P(X \leq 1) = P(X = 0) + P(X = 1) = 56a = 56/64$

(ii) $P(X \leq 1 \mid Y = 2) = P\{X \leq 1, Y = 2\}/P(Y = 2)$

$P(X \leq 1, Y = 2) = P(X = 0, Y = 2) + P(X = 1, Y = 2) = 0 + 4a; \quad P(X \leq 1 \mid Y = 2) = 4a/6a = 2/3.$

(iii) $\quad P(X < 3 \mid Y \leq 4) = P(X < 3, Y \leq 4)/P(Y \leq 4).$

Now $\ P(Y \leq 4) = P(Y = 1) + P(Y = 2) \ldots + P(Y = 4) = 6a + 6a + 11a + 13a = 36a.$

$$P(X < 3, Y \leq 4) = P(X = 0, Y \leq 4) + P(X = 1, Y \leq 4) + P(X = 2, Y \leq 4)$$

$$= 6a + (4a + 4a + 8a + 8a) + (2a + 2a + a + a) = 36a.$$

$\therefore \quad P(X < 3 \mid Y \leq 4) = 36a/36a = 1.$

5*. For the convenience of tabulations we draw joint p.d.f. table taking $a = 0.01 = 1/100$

$y \downarrow$ \ $x \rightarrow$		1 0.1	2 0.2	3 0.4	4 0.2	5 0.1	Total $P_Y \downarrow$
0	0.2	$2a$	$4a$	$8a$	$4a$	$2a$	$20a$
1	0.1	a	$2a$	$4a$	$2a$	a	$10a$
2	0.4	$4a$	$8a$	$16a$	$8a$	$4a$	$40a$
3	0.3	$3a$	$6a$	$12a$	$6a$	$3a$	$30a$
Total $P_X \rightarrow$		$10a$	$20a$	$40a$	$20a$	$10a$	$100a$

$X + Y$	1	2	3	4	5	6	7	8
Prob. :	$2a$	$5a$	$14a$	$19a$	$26a$	$21a$	$10a$	$3a$

XY :	0	1	2	3	4	5	6	8	9	10	12	15
Prob. :	$20a$	a	$6a$	$7a$	$10a$	a	$22a$	$8a$	$12a$	$4a$	$6a$	$3a$

Y/X :	0	$\frac{1}{5}$	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{2}{5}$	$\frac{1}{2}$	$\frac{3}{5}$	$\frac{2}{3}$	$\frac{3}{4}$	1	$\frac{3}{2}$	2	3
Prob. :	$20a$	a	$2a$	$4a$	$4a$	$10a$	$3a$	$16a$	$6a$	$21a$	$6a$	$4a$	$3a$

6*. Let $P(X = 1) = a$, $P(Y = 1) = b$. This immediately provides the marginal totals, $0.8 - b$ and $0.8 - a$. Since X and Y are independent,

$$P(X = 1, Y = 1) = P(X = 1)\, P(Y = 1) \Rightarrow ab = 0.06.$$

$$P(X = 3, Y = 3) = P(X = 3)\, P(Y = 3) \Rightarrow (0.8 - b)\,(0.8 - a) = 0.3.$$

Eliminating b, these equations provide : $a^2 - 0.5a + 0.06 = 0 = (a - 0.2)(a - 0.3)$.
When $a = 0.2$, $b = 0.3$; when $a = 0.3$, $b = 0.2$.

Since marginal totals are known, cell entries are immediately obtainable, by multiplying the row total into column totals. The results are entered in (R.H.S.) table shown in the problem for the case $a = 0.2$ and $b = 0.3$.

7*. Let the first head occur on trial m and the second head occur on trial n, then in n trials two heads and $n - 2$ tails have occurred, so

$$P(X = m, Y = n) = \begin{cases} p^2 q^{n-2}, & m = 1, 2, \ldots, n-1;\ n = 2, 3, \ldots \\ 0, & \text{elsewhere} \qquad p + q = 1 \end{cases} \qquad \ldots(1)$$

By Multi-stage Rule, we get

$$P(X = m) = \sum_{j=m+1}^{\infty} P(X = m, Y = j) = p^2 q^{-1} \sum_{j=m+1}^{\infty} q^{j-1} = p^2 q^{-1} \frac{q^m}{1-q}\, pq^{m-1} \qquad \ldots(2)$$

$$P(Y = n) = \sum_{i=1}^{n-1} P(X = i, Y = n) = (n-1)\, p^2 q^{n-2}. \qquad \ldots(3)$$

$\therefore\ P(X = m \mid Y = n) = P(X = m, Y = n)/P(Y = n) = (n-1)^{-1}$. ...(4)

Thus, if it is known that the second heads occur at the nth trial, the first head must have occurred during the first $(n-1)$ trials. Clearly, all these possibilities are equally likely. Now,

$P(Y = n \mid X = m) = P(X = m, Y = n)/P(X = m) = pq^{n-m-1}$ ...(5)

$$P(Y - X = k \mid X = m) = \frac{P(Y - X = k, X = m)}{P(X = m)} = \frac{P(X = m, Y = m + k)}{P(X = m)} = pq^{k-1} \quad \text{[by (5)]} \quad ...(6)$$

We see from (6), that $Y - X$ is independent of X. Further (2) and (6) reveal that $Y - X$ and X have the same distribution.

8*. The marginal densities are

$$f_1(x) = (1/27)\,\Sigma_y\,(x+2y) = (1/9)\,(x+2), x = 0, 1, 2.$$

$$f_2(x) = (1/27)\,\Sigma_x\,(x+2y) = (1/9)\,(1+2y), y = 0, 1, 2.$$

The conditional distribution of Y for $X = x_0$ is

$$f_{Y\mid x}(Y \mid x_0) = f(x_0, y)/f_1(x_0) = (x_0 + 2y)/(3x_0 + 6).$$

$$\therefore \quad f(y \mid x = 0) = \frac{1}{3}\frac{y}{3},\ f(y \mid x = 1) = \frac{(1+2y)}{9}, \text{ and } f(y \mid x = 2)\frac{(1+y)}{6}.$$

These provides the various values

$f(Y = 0 \mid X = 0) = 0$, $\quad f(Y = 1 \mid X = 0 = \frac{1}{3}$, $\quad f(Y = 2 \mid X = 0) = \frac{2}{3}$

$f(Y = 0 \mid X = 1) = \frac{1}{9}$, $\quad f(Y = 1 \mid X = 1 = \frac{1}{3}$, $\quad f(Y = 2 \mid X = 1) = \frac{5}{9}$,

$f(Y = 0 \mid X = 2) = \frac{1}{6}$, $\quad f(Y = 1 \mid X = 2 = \frac{1}{3}$, $\quad f(Y = 2 \mid X = 2) = \frac{1}{2}$.

Sec. 4-44. Page 168

1*. (i) Let $g(x, y) = f(x)\,f(y)$, since X, Y are i.i.d. Now

$$P(X < Y) = \int_{-\infty}^{\infty} dy \int_{-\infty}^{y} f(x)\,f(y)\,dx = \int_{-\infty}^{\infty} f(y)\,[F(y) - F(-\infty)]\,dy.$$

$$= \int_0^1 t\,dt = \frac{1}{2}. \qquad [t = F(y), f(y)\,dy = dt, F(-\infty) = 0, F(\infty) = 1]$$

(ii) Let $\quad h(x, y, z) = f(x)\,f(y)\,f(z)$, since X, Y, Z are i.i.d. Now

$$P(X < Y < Z) = \int_{-\infty}^{\infty} f(z)\,dx \int_{-\infty}^{z} f(x)\,dx \int_{x}^{z} f(y)\,dy$$

$$= \int_{-\infty}^{\infty} f(z)\,dz \int_{-\infty}^{z} f(x)\,[F(z) - F(x)]\,dx, \qquad \text{[Put } F(z) = u]$$

$$= \int_{-\infty}^{\infty} f(z)\{F^2(z) - \frac{1}{2}F^2(z)\}\ dx, \qquad [f(x) = F'(x)]$$

$$= \frac{1}{2}\int_{-\infty}^{\infty} F^2(z)\,f(z) = \frac{1}{2}\int_0^1 u^2\,du = \frac{1}{3!}.$$

2*. (i) $$p_1 = \frac{P(X>Y, X>Z)}{P(X>Z)} = \frac{P(X=\max(X,Y,Z))}{P(X>Z)} = \frac{\frac{1}{3}}{\frac{1}{2}} = \frac{2}{3}.$$

(ii) $$p_2 = \frac{P(X>Y, X<Z)}{P(X<Z)} = \frac{P(Z>X>Y)}{P(X<Z)} = \frac{\frac{1}{3!}}{\frac{1}{2}} = \frac{1}{3}.$$

(iii) $$p_3 = \frac{P(X>Y, Y>Z)}{P(Y>Z)} = \frac{P(X>Y>Z)}{P(Y>Z)} = \frac{\frac{1}{3!}}{\frac{1}{2}} = \frac{2}{3}.$$

(iv) $$p_4 = \frac{P(X>Y, Y<Z)}{P(Y<Z)} = \frac{P(Y=\min(X,Y,Z))}{P(Y<Z)} = \frac{\frac{1}{3}}{\frac{1}{2}} = \frac{2}{3}.$$

Note that $P(Y<Z) = P(Y>Z) = \frac{1}{2}$ (by symmetry), $P\{X=Y\} = 0$, etc.

3*. Since $f(x, y) = f_1(x) \cdot f_2(y), f_1(x) = \lambda e^{-\lambda x}, x > 0 ; f_2(y) = \lambda e^{-\lambda y}, y > 0$, it follows that X and Y are independent Expo-variates.

(a) $P(X<k) = P(0<x\le k) = \int_0^k \lambda e^{-\lambda x}\, dx = 1 - e^{\lambda k}$.

(b) $$P(X > kY) = \int_{ky}^{\infty} dx \int_0^{\infty} f(x, y)\, dy = \int_0^{\infty} \lambda e^{-\lambda y}\, dy \int_{ky}^{\infty} \lambda e^{-\lambda x}\, dx = \lambda \int_0^{\infty} y^{1-1} e^{-\lambda(1+k)y}\, dy$$

$$= \lambda \cdot \frac{\Gamma(1)}{\lambda(1+k)} = \frac{1}{1+k}.$$

Note by symmetry : $P(X > Y) = P(X<Y) = \frac{1}{2}$, $\quad [P(X>Y) + P(X<Y) = \frac{1}{2}$.

(c) $P(\{X/Y \le k\} = P\{X \le kY\} = 1 - P\{X > kY\} = 1 - 1/(1 + k) = k/(1 + k)$. [by (b)]

Note. If $X/Y = Z$, then $F_Z(k) = k/(1 + k)$, $\quad f'(z) = F'(z) = (1 + z)^{-2}$, $\quad 0 < z < \infty$.

(d) $P(X<Y \mid X<2Y) = P(X<Y,\ X<2Y)/P(X<2Y) = P(X<Y)/P(X<2Y) = \frac{3}{4}$. [$k = 1, 2$, in (c)]

(e) Put $x + y = u$, $y = uv$, so that $dx\, dy = |J|\, du\, dv$;

$\{1 < x + y < 2, x \ge 0, y \ge 0\} \Rightarrow P\{1 < u < 2, 0 < v < 1\}$

$$\therefore \quad P\{1 < X + Y < 2\} = \iint_E \lambda^2 e^{-\lambda(x+y)} dx\, dy = \int_0^1 dv. \int_1^2 \lambda^2 u e^{-\lambda u}\, du$$

$$= \left[(1+\lambda u) e^{-\lambda u}\right]_2^1 = [(1+\lambda) - (1-2\lambda) e^{-\lambda}] e^{-\lambda}$$

$$P\{0 < X + Y < 1\} = \int_0^1 dv \int_0^1 \lambda^2 u e^{-\lambda u}\, du = \left[(1+\lambda u) e^{-\lambda u}\right]_1^0 = 1 - (1+\lambda) e^{-\lambda}.$$

(f) $$p = P(X + Y < m) = P(0 < X + Y < m) = \int_0^1 du \int_0^m u e^{-u} du = \left[(1+u) e^{-u}\right]_m^0 = 1 - (1+m) e^{-u}.$$

When $p = \frac{1}{2}$, this gives $e^m = 2(m + 1)$, with approximate evaluation $m = 1.7$.

(g) Since X and Y are indep. $P(0 < X < 1 \mid Y = 2) = P(0 < X < 1) = 1 - e^{-\lambda}$. [by (a) with $k = 1$].

4*. It is given that y ranges from 0 to x. In order to find the upper limit of x, we suppose it is $k > 0$ and use normalization $p(\Omega) = 0$. Thus

$$1 = \int_0^k \int_0^x f(x, y)\, dx\, dy = 2\int_0^k dx \int_0^x dy = k^2$$

Since $k > 0$, $k^2 = 1$ provides $k = 1$. Thus, $0 < y < x < 1$. Now we find marginal densities.

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y)\, dy = \int_0^x 2dy = 2x, 0 \le x \le 1$$

$$f_Y(y) = \int_{-\infty}^{\infty} f(x, y)\, dx = \int_y^1 2dx = 2(1-y),\ 0 < y < 1.$$

Since $f_1(x)\, f_2(y) \neq f(x, y)$, X and Y are not independent. We now find *Conditional Density Functions* :

$$f_{Y\mid x}(y \mid x) = \frac{f(x, y)}{f_1(x)} = \frac{2}{2x} = \frac{1}{x},\quad 0 < x < 1,\ 0 < y < x.$$

$$f_{X\mid y}(x \mid y) = \frac{f(x, y)}{f_2(y)} = \frac{2}{2(1-y)} = \frac{1}{1-y},\quad 0 < y < x < 1.$$

Try : $P\left\{Y = \frac{1}{2} \mid X = \frac{1}{2}\right\}, P\left\{X \ge \frac{1}{3} \mid Y = \frac{2}{3}\right\}$

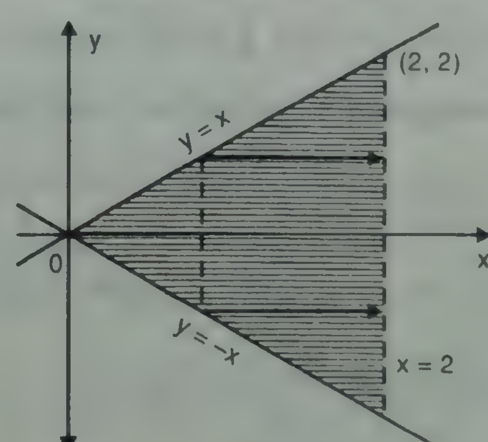
5*. We determine k by normalization $P(\Omega) = 1$. By Even-Odd integrand properties

$$1 = k\int_0^2 dx \int_{-x}^{x} (x^2 - xy)\, dy = 2k \int_0^2 x^2\, dx.$$

$$\int_0^x y = 2k \int_0^2 x^3\, dx = 8k. \text{ Thus } k = \frac{1}{8}.$$

$$\therefore\ f(x, y) = \frac{1}{8}, (x^2 - xy), \qquad 0 < x < 2,\ -x < y < x.$$

$$f_X(x) = \int_{-x}^{x} \frac{1}{8}(x^2 - xy)\, dy = \frac{1}{4}x^2, 0 < x < 2$$



where properties of odd-even integrands are utilized.

$$f_Y(y) = \begin{cases} \frac{1}{8}\int_y^2 (x^2 - xy)\, dx, & 0 \le y \le 2 \\ \frac{1}{8}\int_{-y}^{2} (x^2 - xy)\, dx, & -2 \le y \le 0 \end{cases} \quad \text{i.e.} \quad F_Y(y) = \begin{cases} \frac{1}{3} + \frac{1}{8}\left(\frac{1}{6}y^3 - 2y\right), & 0 \le y \le 2 \\ \frac{1}{3} + \frac{1}{8}\left(\frac{5}{6}y^3 - 2y\right), & -2 \le y \le 0 \end{cases}$$

6*. We check that $k = 2$ by normalization :

$$1 = \int_0^\infty \int_0^x f(x, y)\, dx\, dy = k\int_0^\infty e^{-x}\, dx \int_0^x e^{-y}\, dy = k\int_0^\infty e^{-x}(1 - e^{-x})\, dx = k\left(1 - \tfrac{1}{2}\right) = \tfrac{1}{2}k.$$

$$f_Y(y) = \int_y^\infty 2e^{-(x+y)}\, dx = 2e^{-y}\int_y^\infty e^{-x}\, dx = 2e^{-2y},\ 0 \le y < \infty.$$

$$F_X(x) = \int_0^x 2e^{-(x+y)}\,dy = 2e^{-x}(1-e^{-x}) = 2e^{-x} - 2e^{-2x}, 0 \le x < \infty.$$

$$f_{(x\,|\,y)} = \frac{f(x, y)}{f_Y(y)} = \frac{2e^{-x}.e^{-y}}{2e^{-2y}} = e^{-(x-y)}, y < x < \infty.$$

$$f_{(y\,|\,x)} = \frac{f(x, y)}{f_X(y)} = \frac{2e^{-x}.e^{-y}}{2e^{-x}(1-e^{-x})} = \frac{e^{-y}}{(1-e^{-x})},\quad 0 < y < x$$

$$P(Y \ge 3) = \int_3^\infty f_2(y)\,dy = 2\int_3^\infty e^{-2y}\,dy = e^{-6}.$$

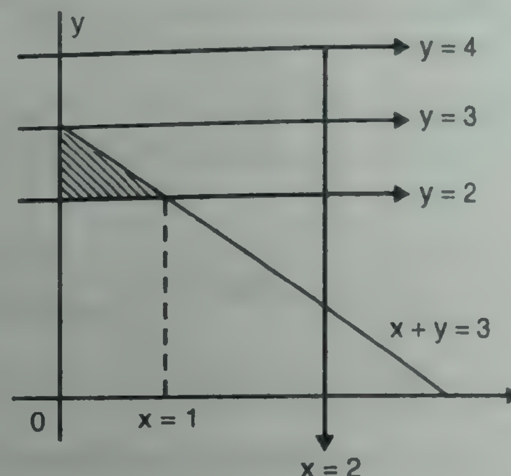
Since $f(x, y) \ne f_X(x)\, f_Y(y)$, it follows that X and Y are not independent.

7*. $P(X < 1, Y < 3) = \int_0^1 \int_2^3 \frac{1}{8}(6 - x - y)\,dx\,dy$

$$= \frac{1}{8}\int_0^1 dx\left[(16 - x)\,y - \frac{y^2}{3}\right]_2^3 = \frac{1}{8}\int_0^1 \left(\frac{7}{2} - x\right)dx = \frac{3}{8}.$$

$$P(Y < 3) = \frac{1}{8}\int_0^2 (6 - x - y)\,dy = \frac{1}{8}\int_0^2 \left(\frac{7}{2} - x\right)dx = \frac{5}{8}.$$

$$\therefore \quad P(X < 1 \mid Y < 3) = \frac{P(X < 1, Y < 3)}{P(Y < 3)} = \frac{\frac{3}{8}}{\frac{5}{8}} = \frac{3}{5}.$$



$$P(X + Y < 3) = \int_0^1 dx \int_2^{3-x} \frac{1}{8}(6 - x - y)\,dy = \frac{1}{8}\int_0^1 \frac{1}{2}(x^2 - 8x + 7)\,dx = \frac{5}{24}.$$

Marginal Distributions and Conditional Distributions :

$$f_1(x) = \int_2^4 \frac{1}{8}(6 - x - y)\,dy = \frac{1}{4}(3 - x),\ 0 < x < 2;\ f_1(x) = 0,\ x \le 0 \text{ or } x \ge 2.$$

$$f_2(y) = \int_0^2 \frac{1}{8}(6 - x - y)\,dx = \frac{1}{4}(5 - y),\ \ 2 < y < 4;\ \ f_2(y) = 0, \text{ otherwise}$$

$$f_3(x \mid y) = \frac{f(x, y)}{f_2(y)} = \frac{6 - x - y}{2(5 - y)},\quad 2 < y < 4,\quad 0 < x < 2.$$

$$f_4(y \mid x) = \frac{f(x, y)}{f_1(x)} = \frac{6 - x - y}{2(3 - x)},\quad 2 < y < 4,\quad 0 < x < 2.$$

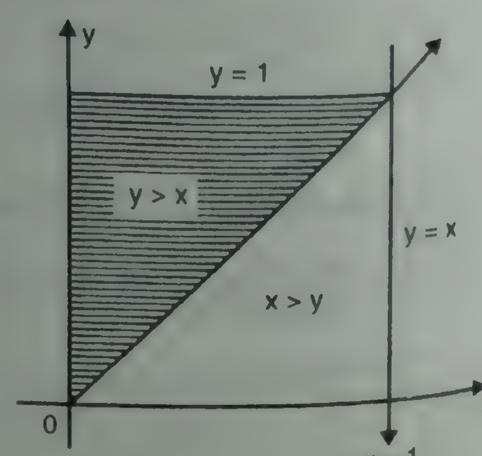
8*. We determine k by normalization $P(S) = 1$. Now

$$f_X(x) = \int_x^1 f(x, y)\,dy = \frac{k}{2}x\left[y^2\right]_x^1$$

$$= \frac{1}{2}kx\,(1 - x^2), 0 < x < 1.$$

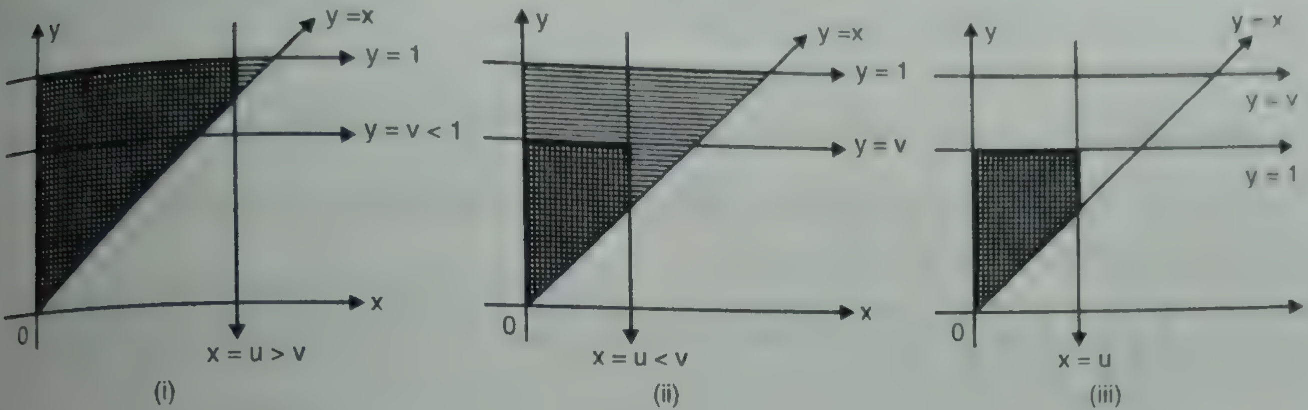
So $\quad 1 = \int_0^1 f_X(x)\,dx = \frac{1}{2}k\left(\frac{1}{2} - \frac{1}{4}\right) \Rightarrow k = 8.$

Thus, $\quad f_1(x) = 4x(1 - x^2), \quad 0 < x < 1.$



$$f_Y(y) = \int_0^y f(x, y)\, dx = 4y\left[x^2\right]_0^y = 4y^3, \quad 0 < y < 1.$$

$$f_1(x \mid y) = \frac{f(x, y)}{f_Y(y)} = \frac{8xy}{4y^3} = \frac{2x}{y^2}, \quad 0 < x < y, \quad 0 < y < 1.$$



$$f_2(y \mid x) = \frac{f(x, y)}{f_X(x)} = \frac{8xy}{4x(1-x^2)} = \frac{2y}{1-x^2}, \quad x < y < 1, \ 0 < x < 1.$$

Since $f(x, y) \neq f_X(x)\, f_Y(y)$, X and Y are not independent. This also follows trivially because the effective record space is not rectangular.

(ii) For distribution functions, we consider the various situations :

We require $P(X \le u, Y \le v)$ and hence the various cases of the positions of lines $x = u$ and $y = v$. See above Fig.

$$F(u, v) = \begin{cases} 0, & u < 0, v < 0 \quad \text{(Space } \varnothing) \\ \int_{y=0}^{v}\int_{x=0}^{y} 8xy\, dx\, dy, & u \ge v, 0 \le v \le 1, \text{ Fig. (i)} \\ \int_{x=0}^{u}\int_{y=0}^{v} 8xy\, dx\, dy, & v \ge u, 0 \le v \le 1, \text{ Fig. (ii)} \\ \int_{x=0}^{u}\int_{x=0}^{y} 8xy\, dx\, dy, & v \ge u, 0 \le u \le 1, \text{ Fig. (iii)} \\ 1, & u \ge v \ge 1 \text{ or } v \ge u \ge 1, \text{ (Space } \Omega) \end{cases}$$

Evaluating the above integrals, we get

$$F(u, v) = \begin{cases} 0, & u < 0, & v < 0 \\ v^4, & u \ge v, & 0 \le v \le 1 \\ 2u^2v^2 - u^4, & v \ge u, & 0 \le v \le 1 \\ 2u^2 - u^4, & v \ge u, & 0 \le u \le 1 \\ 1, & u \ge v \ge 1 \text{ or } & v \ge u \ge 1. \end{cases}$$

(iii) $P\left(Y < \frac{1}{4}\right) = \int_0^{1/4} 4y^3\, dy = \frac{1}{256}$

$$P\left(X < \frac{1}{2}, Y < \frac{1}{4}\right) = \int_{y=0}^{1/4}\int_{x=0}^{y} f(x, y)\, dx\, dy = 8\int_0^{1/4} y\, dy \int_0^y x\, dx = \frac{1}{256}. \quad \text{[Draw Fig.]}$$

$$\therefore \quad P\left(X < \frac{1}{2} \mid Y < \frac{1}{4}\right) = \frac{P(X < \frac{1}{2}, Y < \frac{1}{4})}{P(Y < \frac{1}{4})} = \frac{1/256}{1/256} = 1.$$

9*. (a) $f_X(x) = \int_0^1 f(x, y)\,dy = \frac{\pi^2}{8}\int_0^1 \sin\left[\frac{1}{2}\pi(x+y)\right]dy = \frac{1}{4}\pi\left[\cos\frac{1}{2}\pi(x+y)\right]_1^0$

$$= \frac{1}{4}\pi\left[\cos\frac{1}{2}\pi x + \sin\frac{1}{2}\pi x\right],\ 0 < x < 1.$$

By symmetry of $f(x, y)$ we observe that, $f_Y(y) = \frac{1}{4}\pi\left[\cos\left(\frac{1}{2}\pi y\right) + \sin\left(\frac{1}{2}\pi y\right)\right],\ 0 < y < 1.$

Since $f(x, y) \neq f_X(x)\,f_Y(y)$; X and Y are not independent although the Record Space is a rectangular region : $0 \le x \le 1,\ 0 \le y \le 1.$

(c) $P\left(Y \le \frac{1}{2}X\right) = \int_0^1 dx \int_0^{x/2} \frac{\pi^2}{8}\sin\left(\frac{\pi x}{2} + \frac{\pi y}{2}\right)dy = \frac{\pi}{4}\int_0^1\left[\cos\left(\frac{\pi x}{2}\right) - \cos\left(\frac{3\pi x}{4}\right)\right]dx = \frac{1}{2}\left(1 - \frac{\sqrt{2}}{3}\right).$

10*. Recall : $\sum_{r=1}^{\infty} z^r = \frac{z}{1-z}, |z| < 1\ \sum_{r=1}^{\infty} r\,.\,z^{r-1} = \frac{1}{(1-z)^2}.$ [by Differentiation] ...(1)

Now : $f(x, y) = f_1(x)\,f_2(y \mid x) = qx(py)^{x-1}, x = 1, 2, 3, \ldots, 0 \le y \le 1$ (Joint density) ...(2)

The unconditional density of Y, say $g(y)$, is obtained from (2) by summing out x

$\therefore\ g(y) = \sum_{x=1}^{\infty} f(x, y) = q\sum_{x=1}^{\infty} x(py)^{x-1} = \frac{q}{(1-py)^2},$ [by (1) b] $0 \le y \le 1,$ ...(3)

So $f_3(x \mid y) = f(x, y)/g(y) = x(py)^{x-1}\,.\,(1-py)^2,\ 0 \le y \le 1,\quad x = 1, 2, \ldots$ ...(4)

(i) $P(Y > \frac{1}{2} \mid X = k) = \int_{1/2}^1 f_2(y \mid X = k)\,dy = \int_{1/2}^1 k\,.\,y^{k-1}\,dy = 1 - (\frac{1}{2})^k.$

(ii) $P\{X = k \mid Y < \frac{1}{2}\} = P\{X = k, Y < \frac{1}{2}\}/P(Y < \frac{1}{2})$ ...(5)

Now, $P\left(Y < \frac{1}{2}\right) = \int_0^{1/2} q(1-py)^{-2}\,dy = \frac{q}{2-p}.$

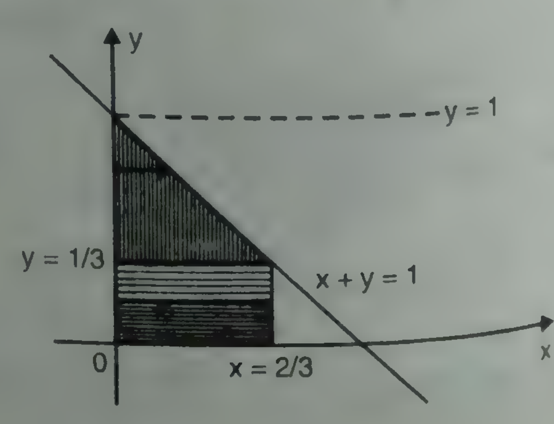
$P\left(X = k, Y < \frac{1}{2}\right) = kq\int_0^{1/2}(py)^{k-1}\,dy = qp^{k-1}\left(\frac{1}{2}\right)^k.$ [using $x = k$ in Eq. (2)]

Substituting these values into (5) we get, $P\{X = k \mid Y < \frac{1}{2}\} = (2-p)\,p^{k-1}\,(\frac{1}{2})^k.$

11*. The sample space is exhibited in the diagram. We find the joint p.d.f. and marginal density of Y.

$f(x, y) = f(y \mid x)\,f_1(x) = [2y/(1-x)^2]\,[81(1-x)^2/26]$

$= (81/13)\,y,\ 0 < y < 1 - x,\ 0 < x < \frac{2}{3}.$



$$f_Y(y) = \begin{cases} \int_0^{2/3}\left(\frac{81}{13}\right)y\,dx, & 0 < y < \frac{1}{3} \\ \int_0^{1-y}\left(\frac{81}{13}\right)y\,dx, & \frac{1}{3} < y < 1 \end{cases}$$

$$\therefore \quad f_Y(y) = \begin{cases} \left(\dfrac{54}{13}\right) y, & 0 < y < \dfrac{1}{3} \\ \left(\dfrac{81}{13}\right) y(1-y), & \dfrac{1}{3} < y < 1 \end{cases}.$$

The conditional density of X given $Y = y$, i.e. $g(x \mid y) = f(x, y)/f_Y(y)$ is given by

$$g(x \mid y) = \begin{cases} \left(\dfrac{81y}{13}\right) \Big/ \left(\dfrac{54y}{13}\right) = \dfrac{3}{2}, & 0 < y < \dfrac{1}{3}, \quad 0 < x < \dfrac{2}{3} & \text{...(i)} \\ \left(\dfrac{81y}{13}\right) \Big/ \left[\dfrac{81y(1-y)}{13}\right] = (1-y)^{-1}, & \dfrac{1}{3} < y < 1, \quad 0 < x < 1 - y & \text{...(ii)} \end{cases}$$

When $Y = 30 \text{ min} = \dfrac{1}{2}$ hr, we get from (ii), $g\left(x \mid \tfrac{1}{2}\right) = 2, 0 < x < \tfrac{1}{2}$.

$$\therefore \quad P\left(X \le \frac{1}{4}\right) = \int_0^{1/4} 2\, dx = \frac{1}{2} \qquad \left(\because \ 15 \text{ min.} = \frac{1}{4} \text{ hr.}\right)$$

When $Y = 50$ min. $= \left(\tfrac{5}{6}\right)$ hr, then the event of leaving bed later than 6.20 a.m., i.e. $\{X \ge \tfrac{1}{3}\}$ is impossible, because $x \ge \tfrac{1}{3}$, $y = \tfrac{5}{6}$ gives $x + y = \tfrac{7}{6} > 1$. Hence $P(X \ge \tfrac{1}{3}) = 0$.

Sec. 4-55. Page 178

1*. By convolution of independent variates

$$f_Z(z) = \int_{-\infty}^{\infty} f_1(x) f_2(z - x)\, dx = \int_{z-1}^{z} f_1(x)\, dx = F_1(z) - F_1(z-1),$$

because $f_2(z - x) = 1$, $0 \le z - x \le 1$, as $Y \sim U(0, 1)$. We note that

$$0 \le y \le 1 \Rightarrow 0 \le z - x \le 1 \quad \Rightarrow z - 1 \le x \le z.$$

2*. Here $f_1(x) = \lambda e^{-\lambda x}$, $P(X \le k) = \int_0^k \lambda e^{-\lambda x}\, dx = 1 - e^{-\lambda k}$ ...(1)

Now $\quad F_U(z) = P\{U \le u\} = 1 - P\{\min(X, Y) > u\} = 1 - P(X \ge u) P(Y \ge u) = 1 - e^{-\lambda u} \cdot e^{-\lambda u} = 1 - e^{-2\lambda u}$

$\therefore \quad f_U^{(u)} = 2\lambda e^{-2\lambda u}$, $u > 0$, $f_U^{(u)} = 0$, elsewhere. $\qquad [F'(u) = f(u)]$

(ii) $\quad F_V^{(v)} = P\{V \le v\} = P\{\max(X, Y) \le v\} = P\{X \le v, Y \le v\}$

$\qquad = P(X \le v) \cdot P(Y \le v) = (1 - e^{-\lambda v})^2 \qquad$ [by indep. and by (1)]

$\therefore \quad f_V^{(v)} = 2\lambda e^{-\lambda v}(1 - e^{-\lambda v})$, $v > 0$; $f_V^{(v)} = 0$, elsewhere.

Sec. 4-62. Page 182

1*. (a) Here, $4xye^{-x^2 - y^2} = (2xe^{-x^2})(ye^{-y^2}) \Rightarrow f(x, y) = f_1(x) \cdot f_2(y)$.

This shows that X and Y are independent distributed with densities :

$$f_1(x) = 2xe^{-x^2}, x > 0; \ f_2(y) = 2ye^{-y^2}, \ y > 0.$$

By independence $f(x \mid Y = y) = f_X(x) = 2xe^{-x^2}, x \ge 0$.

(b) We change to polar coordinates : $x = u \cos\theta$, $y = u \sin\theta$, $\left|\dfrac{\partial(x, y)}{\partial(u, \theta)}\right| = u.$

Also $x \geq 0$, $y \geq 0$ (the positive quadrant) transform to $0 \leq u < \infty$, $0 \leq \theta \leq \frac{1}{2}\pi$. The joint p.d.f. is thus

$$g(u, \theta) = f(x, y)|J| = 4u^2 \sin\theta \cos\theta\, e^{-u^2} |u| \qquad ...(1)$$

To find the density of u, we integrate out the unwanted variate θ from (1). This gives

$$\therefore \quad g(u) = 4u^3 e^{-u^2} \int_0^{\pi/2} \sin\theta \cos\theta\, d\theta = 2u^3 e^{-u^2}, u \geq 0.$$

2*. Let $z = xy$, $w = x$; then $\left|\dfrac{\partial(z, w)}{\partial(x, y)}\right| = |-x| = |w|$; so $dx\,dy = dw\,dz \mid w \mid$. The joint probability differential of X, Y is

$$dF_{X, Y}(x, y) = (\pi\sqrt{1-x^2})^{-1} . y e^{-y^2/2}\, dx dy, \ -1 < x < 1, \ 0 < y < \infty \qquad ...(1)$$

$$\therefore \quad F_{Z, W}(z, w) = \frac{dz dw}{\pi\sqrt{1-w^2}}\left(\frac{z}{w}\right)\frac{\exp(-z^2/2w^2)}{|w|}, \quad -\infty < z < \infty, \ 0 < w < 1 \ ...(2)$$

Note that $w < 0$, since for $z > 0$, $f(z, w) < 0$ when $w < 0$, which is not acceptable. To find p.d.f. Z, we integrate out the unwanted W. Thus

$$dF_Z(z) = \frac{zdz}{\pi}\int_0^1 \frac{\exp(-z^2/2w^2)}{w^2\sqrt{1-w^2}}\, dw = \frac{ze^{-z^2/2} dz}{2\pi}\int_0^\infty \exp\left(-\frac{z^2 t}{2}\right) . t^{(1/2)-1}\, dt.$$

$$f_Z(z) = \frac{ze^{-z^2/2}}{\pi}\frac{\Gamma(1/2)}{\sqrt{1/2z^2}} = \frac{e^{z^2/2}}{\sqrt{2}\,\pi}, \quad -\infty < z < \infty.$$

3*. (i) Here $f(x, y) = x^{a-1} y^{b-1} (1 - x - y)^{c-1} / B(a, b, c), x \geq 0, y \geq 0, x + y \leq 1.$ ...(1)

Now $x = z$, $y = t(1 - z)$, $| \partial(x, y)/\partial(z, t) | = 1 - z$, $dxdy = (1 - z)\, dz\, dt$.

$1 - x - y = (1 - t)(1 - z)$; $x + y \leq 1 \Rightarrow 0 \leq t, z \leq 1$.

$$g(z, t) = f(x, y)|J| = z^{a-1} . [t(1-z)]^{b-1} [(1-t)(1-z)]^{c-1} . (1-z) / B(a, b, c)$$

$$= \frac{z^{a-1}(1-z)^{b+c-1}}{B(a, b+c)} . \frac{t^{b-1}(1-t)^{c-1}}{B(b, c)}, \quad 0 \leq z, \ t \leq 1.$$

It follows that $Z \sim B_I(a, b + c)$ and $T \sim B_I(b, c)$ are independent.

(ii) Let $u + v + w = s$ then set $u = ps$, $v = s$, $w = (1 - p - q)\, s$. Thus

$$\left|\frac{\partial(u, v, w)}{\partial(p, q, s)}\right| = s^2. \text{ Now } u \geq 0, v \geq 0, w \geq 0 \Rightarrow p \geq 0, q \geq 0, p + q \leq 1.$$

The joint density of indep. variates U, V, W is

$$f(u, v, w) = \frac{e^{-u} u^{l-1}}{\Gamma(l)} . \frac{e^{-v} v^{l-1}}{\Gamma(m)} . \frac{e^{-w} u^{w-1}}{\Gamma(n)}$$

$\therefore \quad g(p, q, s) = f(u, v, w) \mid J \mid = \dfrac{e^{-s}(ps)^{\ell-1}(qs)^{m-1}[(1-p-q)s]^{n-1} . s^2}{\Gamma(\ell)\Gamma(m)\Gamma(n)}$

We integrate *out* s, $(0 < s < \infty)$ to obtain

$$g(p, q) = \frac{p^{\ell-1}q^{m-1}(1-p-q)^{n-1}}{\Gamma(\ell)\Gamma(m)\Gamma(n)} \int_0^\infty e^{-s} S^{(\Sigma\ell)-1} . ds = \frac{p^{\ell-1}q^{m-1}(1-p-q)^{n-1}\Gamma(\Sigma\ell)}{\Gamma(\ell)\Gamma(m)\Gamma(n)}$$

$$= \frac{p^{\ell-1}q^{m-1}(1-p-q)^{n-1}}{B(\ell, m, n)}; \quad p \geq 0, \ q \geq 0, \ p+q \leq 1.$$

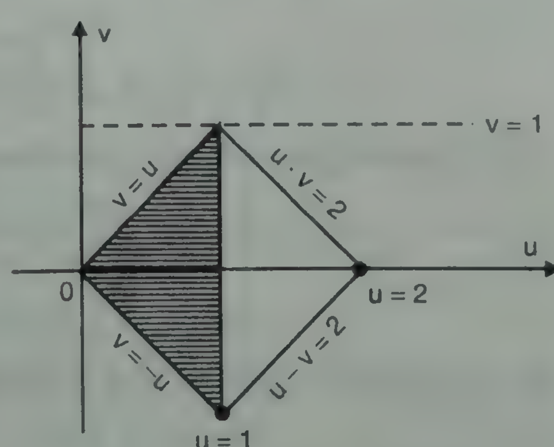
4*. As usual, $f_1(x) = k\int_0^{1-x}(x+y)\,dy = k\left[x(1-x)+\frac{1}{2}(1-x)^2\right] = \frac{1}{2}k(1-x^2).$

Since $1 = \int_0^1 f_1(x)\,dx = \frac{1}{2}k\left(1-\frac{1}{3}\right) = \frac{k}{3}$, hence $k = 3$ and $f_1(x) = \dfrac{3(1-x^2)}{2}$, $0 < x < 1$.

Joint density. Let $u = x + y$, $v = x - y$, so that $x = \frac{1}{2}(u+v)$, $y = \frac{1}{2}(u-v)$, $\left|\dfrac{\partial(x, y)}{\partial(u, v)}\right| = \dfrac{1}{2}$;

$$0 < x < 1 \Rightarrow 0 < u + v < 2, \quad 0 < y < 1 \Rightarrow 0 < u - v < 2$$
$$0 < x + y < 1 \Rightarrow 0 < u < 1, x > 0 \Rightarrow u + v > 0, \quad y > 0 \Rightarrow u > v$$

The region of integration is shown in the Figure.



The joint p.d.f. of $U.V$ is $g(u, v)$ and is given by

$g(u, v) = f(x, y) \mid J \mid = 3u/u \quad (u, v) \in D$

The marginal density of U is

$g_1(u) = \int_{-u}^{u}\left(\frac{3}{2}\right)u\,dv = 3u\int_0^u dv = 3u^2, \quad 0 < u < 1.$

The marginal density of V is

$g_2(v) = \int_v^1\left(\frac{3}{2}\right)u\,du, 0 < v < 1; g_2(v) = \int_{(-u)}^1 u\,du, \ -1 < v < 0$

i.e. $\qquad g_2(v) =$

$\left(\frac{3}{4}\right)(1-v^2), 0 < v < 1; g_2(v) = \left(\frac{3}{4}\right)(1-v^2), -1 < v < 0$

Combining : $\qquad g_2(v) = \left(\frac{3}{4}\right)(1-v^2), \ -1 < v < 1.$

Sec. 4-70. Page 186

1*. By Normalization, $\Sigma p_{ij} = 96a = 1$ so that $a = 1/96$.

$P(x_3, y_4) = 39\,a \neq 64a \,.\, 60a = P(x_3)\,P(y_4) \Rightarrow X$ and Y are not independent.

Note. For the pairs (x_1, y_1), (x_1, y_2), (x_1, y_3), (x_2, y_1), (x_2, y_2), (x_3, y_1), (x_3, y_2), it can be verified that $P(x_i, y_j) = P(x_i)\,P(y_j)$.

So don't haste to conclude independence, rather try to find some counter example as we have done or exhaust all possibilities.

2*. $P(XY = 1) = P\{(X = 1, Y = 1) \uplus (X = -1, Y = -1)\} = P(X = 1, Y = 1) + P(X = -1, Y = -1)$

$$= P(X = 1)\, P(Y = 1) + P(X = -1)\, P(Y = -1) = (\tfrac{1}{4}) + (\tfrac{1}{4}) = \tfrac{1}{2}$$

$$P(XY = -1) = P[(X = 1, Y = -1) \uplus X = -1, Y = 1)]$$

$$= P(X = 1, Y = -1) + P(X = -1, Y = 1) = (\tfrac{1}{4}) + (\tfrac{1}{4}) = \tfrac{1}{2}.$$

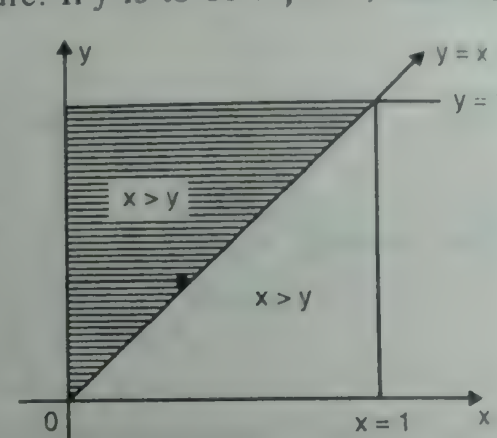
Thus $P(Z = \pm 1) = \frac{1}{2}$, where $Z = XY$. We check independent of X and Z.

$P(X = 1, Z = 1) = P(X = 1, Y = 1) = P(X = 1)\, P(Y = 1) = \frac{1}{4}; P(X=1, Z=-1) = P(X = 1, Y = -1) = \frac{1}{4}$

$P(X = -1, Z = 1) = P(X = -1, Y = -1) = \frac{1}{4}; P(X = -1, Z = -1) = P(X = -1, Y = 1) = \frac{1}{4}$.

Thus $P\{(X = \pm 1) \cap (Z = \pm 1)\} = P(X = \pm 1)\, P(Z = \pm 1) \Rightarrow X$ and $Z = XY$ are independent.

3*. The region of integration is that shown in the figure. If f is to be a p.d.f., then we must have, $k > 0$, and



$$1 = \int_0^1 dy \int_0^y (y - x)^\beta\, dx. \qquad ..(1)$$

(i) The x-integral is convergent only if $\beta + 1 \neq 0$. Further, $\beta + 1$ must be positive i.e. $\beta > -1$, otherwise $f(x, y) < 0$. Hence, when $\beta > -1$, we can choose k so as to make f a p.d.f.

(ii) Assuming $\beta > -1$, we complete integration in (1) to get

$$\frac{1}{k} = \int_0^1 \frac{y^{\beta+1}}{\beta+1}\, dy = \frac{1}{(\beta+1)(\beta+2)} \Rightarrow k = (\beta+1)(\beta+2).$$

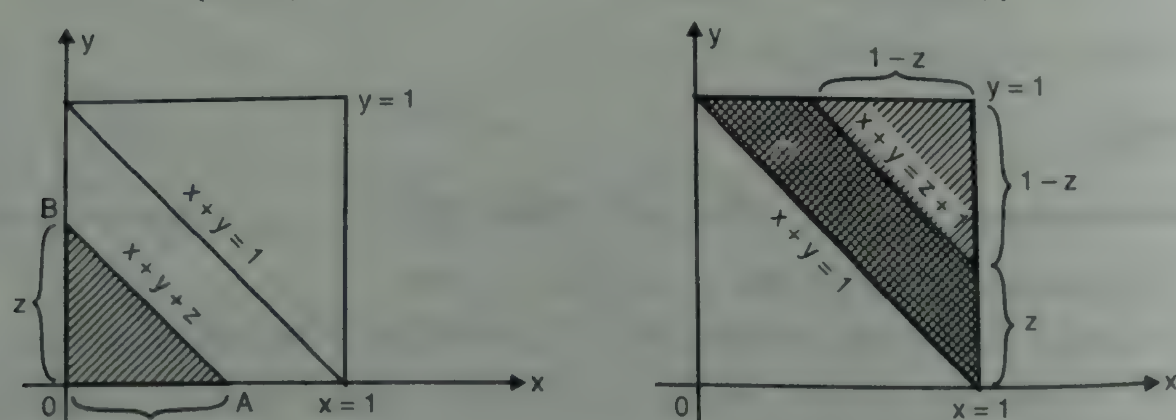
(iii) We now find marginal densities of X and Y,

$$f_X(x) = k\int_x^1 (y - x)^\beta\, dy = \frac{k}{\beta+1}(1 - x)^{\beta+1} = (\beta+2)(1 - x)^{\beta+1}, \quad 0 < x < 1$$

$$f_Y(y) = k\int_x^1 (y - x)^\beta\, dx = \left[(\beta+2)(y - x)^{\beta+1}\right]_y^0 = (\beta+2)\, y^{\beta+1}, \quad 0 < y < 1.$$

4*. The given ranges of $x + y$ provide that $0 \le z \le 1$. Now

$$\{Z \le z\} = \{(Z \le z : X + Y \le 1) \uplus (Z \le z : X + Y \ge 1)\}$$



(i) (ii)

$$\therefore \quad P(Z \le z) = P\{Z \le z : X + Y \le 1\} + P\{Z \le z : X + Y \ge 1\}$$

$$= P\{X + Y \le z\,;\, X + Y \le 1\} + P\{X + Y - 1 \le z\,;\, X + Y \ge 1\} \quad ...(1)$$

Now $P\{X+Y \le z\,;\, X+Y \le 1\} = \frac{1}{2}z^2$ [shaded area in Fig. (i)]

$P\{X+Y \le z+1 : X+Y \ge 1\} = \frac{1}{2} - \frac{1}{2}(1-z)^2$, [shaded area in Fig. (ii)]

Substitutions into (1) provide $P(Z \le z) = \frac{1}{2}z^2 + \frac{1}{2} - \frac{1}{2}(1-z)^2 = z, 0 \le z \le 1$.

$\therefore \quad g_z(z) = F'_z(z) = 1,\ 0 \le z \le 1$.

5*. Let $U = X$, $V = X + Y$; put $x = u$, $y = v - u$; then $|\partial(x, y)/\partial(u, v)| = 1$.

Now $\quad 0 < x < y \Rightarrow 0 < 2u < v < \infty$. The joint p.d.f. of U, V is thus

$$g(u, v) = f(x, y)\,|J| = 2e^{-v},\ 0 < 2u < v < \infty.$$

Marginal distributions

$$g_1(u) = \int_{2u}^{\infty} 2e^{-v}\,du = 2e^{-2u}, u \ge 0; \quad g_2(v) = \int_0^{v/2} 2e^{-v}\,du = ve^{-v},\ v \ge 0.$$

6*. f is everywhere to denote the p.d.f. The arguments involved shall clear the context of the random variables referred to.

$$f(x, y, z) = f(x, y \mid z)\,.\,f(z) = \tfrac{1}{2}[z + (1-z)], (x+y), 0 < x, y < 1, 0 \le z \le 2.$$

(a) To find $f(x, y)$, we integrate **out** z to get

$$f(x, y) = \frac{1}{2}\int_0^2 [z + (1-z)(x+y)]\,dz = 1, 0 < x, y < 1.$$

$$\therefore \quad f_X(x) = \int_0^1 1\,dy = 1\ ;\ f_Y(y) = \int_0^1 1\,dx = 1.$$

Thus $\quad f(x, y) = f_X(y)\,.\,f_X(y) \quad \Rightarrow \quad X$ and Y are independent.

(b) To find $f(x, z)$, we integrate **out** y from $f(x, y, z)$; thus

$$f(x, z) = \tfrac{1}{2}\int [z + (1-z)(x+y)]\,dy = \tfrac{1}{4}[(1+z) + 2(1-z)\,x],\ 0 < x < 1, 0 \le z \le 2.$$

Since $f_X(x) = 1, f_Z(z) = \frac{1}{2}$, we observe that $f(x, z) \ne f_X(x)\,f_Z(z) \Rightarrow X$ and Z are not independent.

(c) Let $X + Y = V$, $X = U$, then $\ x = u, y = v - u \mid \partial(x, y)/\partial(u, v) \mid = 1$.

The joint distribution of U and V is thus $g(u, v) = f(x, y) \mid J \mid$, i.e.

$$g(u, v) = 1\ ;\ 0 < u < 1, \quad 0 < v < 2.$$

(d) Let $W = \max(X, Y)$, so that $0 \le w \le 1$, since max $X = 1$, max $Y = 1$.

$$\therefore \quad P\{W \le w \mid Z = z\} = P\{X \le w, Y \le w \mid Z = z\}$$

$$= \int_0^w \int_0^w f(x, y \mid z)\,dx\,dy = \int_0^w \int_0^w [(z + (1-z)(x+y)]\,dx\,dy$$

$$= zw^2 + (1-z)\,w^3,\ 0 < w < 1.$$

(e) Let $x + y = v$, $x = u$, so that $x = u$, $y = v - u$, $|J| = 1$.

$\therefore\ g[u, v] \mid z] = f[(x, y) \mid z] \mid J \mid = z + (1-z)\,v,\ 0 < u < 1,\ u < v < u + 1,\ 0 \le v \le 2.$

To find the marginal density of V, we integrate **out** u to get

$$f(v \mid z) = \int g(u, v \mid z)\,du$$

Now, if $0 \le v \le 1$, then $0 \le u \le v \le 1$;

if $1 \le v \le 2$; then $u \le 1 \le v \le u + 1 \le 2$ or $v - 1 \le u < 1$.

$$\therefore f(v \mid z) = \begin{cases} \int_{u=0}^{v} [z + (1-z)\,v]\,du, & 0 \le v \le 1 \\ \int_{v-1}^{1} [z + (1-z)\,v]\,du, & 1 \le v \le 2 \end{cases} \quad \text{i.e } f(v \mid z) = \begin{cases} [z + (1-z)\,v]\,v, & 0 \le v \le 1 \\ [z + (1-z)\,v](2 - v), & 1 \le v \le 2 \end{cases}$$

Chapter 5 : Mathematical Expectation

Sec. 5-32. Page 198

1*. (a) $\mathrm{E}(|X|) = \Sigma x \,.\, (1/x) = \Sigma 1 = 1 + 1 + 1 + \ldots$

The sum involved in unbounded; hence $\mathrm{E}(X)$ is not defined.

(b) $$\mathrm{E}\,|X| = \sum_{i=1}^{\infty} |x_i|\, f(x_i) = \sum_{i=1}^{\infty} \frac{3^i}{i} \cdot \frac{2}{3^i} = 2\sum_{i=1}^{\infty}\left(\frac{1}{i}\right).$$

The series $\Sigma(1/i)$ is well-known divergent Harmonic series. Hence, $\Sigma(|X|)$ does not exist.

Note. $\displaystyle\sum_{i=1}^{\infty} \frac{(-1)^{i+13^i}}{i} \frac{2}{3^i} = 2\sum_{i=1}^{\infty}\frac{(-1)^{i+1}}{i}$, is a conditionally convergent series.

(c) $\displaystyle \mathrm{E}(|X|) = \sum_{x=1}^{\infty} \frac{|x|}{x(x+1)} = \sum_{x=1}^{\infty}\left(\frac{1}{x+1}\right)$

This is divergent (harmonic) series, hence $\mathrm{E}(X)$ does not exist.

(d) $\displaystyle \mathrm{E}(|X|) = \int_0^{\infty} x \,.\, \frac{dx}{(x+1)^2}$: This integral does not converge.

(e) $\displaystyle \mathrm{E}(|X|) = \int_{-\infty}^{\infty} |x| \,.\, \frac{dx}{\pi(1+x^2)} = \frac{2}{\pi}\int_0^{\infty} \frac{x\,dx}{1+x^2} = \frac{1}{\pi}[\log(1+x^2)]_0^{\infty} \to \infty.$

Since the integral does not converge absolutely, $\mathrm{E}(X)$ does not exist for the given p.d.f.

2*. Assume X to be continous with an integrating density $f(x)$. If $\mathrm{E}(X)$ exists, then

$$\mathrm{E}(X) = \int_{-\infty}^{\infty} x f(x)\, dx < \infty \qquad \ldots(1)$$

and the integral on the R.H.S. of (1) is absolutely convergent. Hence using $f(x) \geq 0$

$$\int_{-\infty}^{\infty} |x\, f(x)|\, dx = \int_{-\infty}^{\infty} |x|\, f(x)\, dx < \infty \Rightarrow \mathrm{E}(|X|) < \infty.$$

Conversely, let $\mathrm{E}(|X|) < \infty$, then $\int_{-\infty}^{\infty} |x|\, f(x)\, dx < \infty$

This shows that $\int_{-\infty}^{\infty} x f(x)\, dx$ converges absolutely $[f(x) \geq 0]$, i.e. $\mathrm{E}(X)$ certainly exists. Further

$$\left|\int_{-\infty}^{\infty} x f(x)\, dx\right| \leq \int_{-\infty}^{\infty} |x|\, f(x)\, dx \;\Rightarrow\; |\mathrm{E}(X)| \leq \mathrm{E}(|X|).$$

If X is discrete, the proof follows by replacing integration by summation of series.

Note. $|X| - X \geq 0 \Rightarrow \mathrm{E}(|X|) \geq \mathrm{E}(X) \Rightarrow \mathrm{E}(X) \leq \mathrm{E}(|X|) \Rightarrow |\mathrm{E}(X)| \leq \mathrm{E}(|X|).$

3*. (i) $$1 = \mathrm{E}\left(\frac{S_n}{S_n}\right) = \mathrm{E}\left(\frac{X_1 + \ldots + X_n}{S_n}\right) = \mathrm{E}\left(\frac{X_1}{S_n} + \frac{X_2}{S_n} + \ldots + \frac{X_n}{S_n}\right) \qquad [\text{Lin E}]$$

$$= \mathrm{E}\left(\frac{X_1}{S_n}\right) + \mathrm{E}\left(\frac{X_2}{S_n}\right) + \ldots + \left(\frac{X_n}{S_n}\right) = n\mathrm{E}\left(\frac{X_i}{S_n}\right) \qquad [\because X_j \text{ are i.i.d.}]$$

Thus $\quad E(X_i/S_n) = 1/n.$ ...(i)

$E(S_k/S_n) = E[(X_1 + \dots + X_k)/S_n] = E(X_1/S_n) + \dots + E(X_k/S_n) = kE(X_1/S_n) = k/n$ [by Lin E]

(ii) $t = E(S_n \mid S_n = t) = E(X_1 \mid S_n = t) + E(X_2 \mid S_n = t) + \dots + E(X_n \mid S_n = t) = nE(X_1 \mid S_n = t)$ [By Lin E]

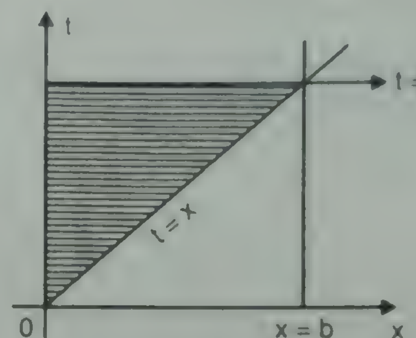
Thus, $E(X_1 \mid S_n = t) = t/n$ by (1)

Warning. Distinguish between the symbols, fractions / and conditioning '|' signs

4*. $E(Y) = \int_{-\infty}^{\infty} y \,.\, f(x)\, dx = \int_0^{\infty} y f(x)\, dx = \int_0^3 (2x+3)\, e^{-x}\, dx + \int_3^{\infty} x e^{-x}\, dx \; [y = y(x) > 0]$

$$= (5 - 11e^{-3}) + 4e^{-3} = 5 - 7e^{-3}.$$

5*. Let $f(x)$ be the p.d.f. of X. Let J denote the given integral then



$$J = \int_0^b [1 - F(x)\, dx] = \int_0^b \left[\int_{t=x}^{b} f(t)\, dt\right] dx = \int_{t=x}^{b} f(t)\, dt \int_{x=0}^{x=t} dx$$

Assuming the change in the order of integration of region OAB.

$\therefore \quad J = \int_0^b t f(t)\, dt = E(X)$ [Kernel definition]

6*. Since variates are i.i.d. and Y is min $\{X_1, \dots, X_m\}$, so

$$P(Y \geq n) = P(X_1 \geq n, \dots, X_m \geq n) = P(X_1 \geq n)\,.\,P(X_2 \geq n) \dots [P(X_m \geq n)]\,[P(X_1 \geq n)]^m = (r_n)^m$$

because $P(X_1 \geq n) = \sum_{k=n}^{\infty} P(X_1 = k) = \sum_{k=n}^{\infty} p_k = r_n.$

Now, for integer-valued variate : $E(Y) = \sum_{n=1}^{\infty} P(Y \geq n) = (r_n)^m$ ...(1)

When $X_j \sim$ gem(p), $P(X_j = x)\; q^{x-1} p$, $x = 1, 2, \dots$, so

$$r_n = P(X_i \geq n) = \sum_{x=n}^{\infty} q^{x-1} p = pq^{n-1}(1 + q + q^2 + \dots) = q^{n-1}.$$

$\therefore \quad E(Y) = \sum_{n=1}^{\infty} (q^m)^{n-1} = 1/(1 - q^m),$ [by (1)]

This shows that $Y = \min\{X_1, \dots, X_n\} \sim$ gem (p'), where $p' = 1 - q^m$.

7*. Here, use of upper-tail probabilities is beneficial. Call D_r = rth draw. Now

$P(X > 1) = 1$, because at least two draws are always required

$$P(X > 2) = P\{\text{Distinct numbers on } D_1, D_2\} = \frac{n}{n} \cdot \frac{n-1}{n} = 1 - \frac{1}{n}.$$

$$P(X > 3) = P\{\text{Distinct Nos. on } D_1, D_2, D_3\} = \frac{n}{n} \cdot \frac{n-1}{n} \cdot \frac{n-2}{n} = \left(1 - \frac{1}{n}\right) \cdot \left(1 - \frac{2}{n}\right),$$

$$P(X > k) = 1 . \left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right) \dots \left(1 - \frac{k-1}{n}\right), \qquad k = 1, 2, 3, \dots$$

$$P\{X=k\}=P(X>k-1)-P(X>k)=\left(1-\frac{1}{n}\right)\ldots\left(1-\frac{k-2}{n}\right)-\left(1-\frac{1}{n}\right)\ldots\left(1-\frac{k-1}{n}\right)$$

$$p_k=\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\ldots\left(1-\frac{k-2}{n}\right)\left(\frac{k-1}{n}\right)\quad k=2,3,\ldots\text{(p.m.f.)}$$

$$\therefore\quad E(X)=\sum_{k=0}^{\infty}P(X>k)=P(X>0)+P(X>1)+P(X>2)+\ldots$$

$$=1+1+\left(1-\frac{1}{n}\right)+\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)+\ldots+\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\ldots\left(1-\frac{n-1}{n}\right)$$

Obviously, $P(X>k)=0$, for $k>n$.

8*. (i) As $P(X=x_i)=P(X\geq x_i)-P(X\geq x_{i+1})$; we rewrite it as $p_i=P_i-P_{i+1}$ (with obvious meaning). We now have

$$E(X)=\Sigma p_i x_i=\Sigma x_i(P_i-P_{i+1})=\Sigma x_i P_i-\Sigma x_i-P_{i+1},\quad [P_{n+1}=0]$$

$$=\sum_{i=1}^{n}x_iP_i-\sum_{i=2}^{n}x_{i-1}P_i=\sum_{i=1}^{n}(x_i-x_{i-1})P_i,\qquad [x_0=0]$$

(ii) We multiply the stated relation : $P(X\geq x_i)\geq P(Y\geq x_i)$ by $(x_i-x_{i-1})>0$ and sum up to get

$$\Sigma(x_i-x_{i-1})\,P(X\geq x_i)\geq\Sigma(x_i-x_{i-1})\,P(Y\geq x_i),i=1,2,\ldots,n$$

$$\therefore\qquad E(X)\geq E(Y)\qquad\text{[by Part (i)]}$$

Relation (1) implies that distribution of Y is shifted to the left of the distribution of X]

9*. Define the indicator X_j by : $X_j=1$, if A_j occurs; $X_j=0$, if $\overline{A}_j$ occurs.

Then, $E(X_j)=P(A_j)$. Now

$$0\leq\left(\sum_{j=1}^{n}a_jX_j\right)^2=\left(\sum_{i=1}^{n}a_iX_i\right)\left(\sum_{j=1}^{n}a_jX_j\right)=\sum_{i=1}^{n}\sum_{j=1}^{n}a_ia_jX_iX_j.$$

Nothing $E(X_iX_j)=P(A_iA_j)$ and taking expected values in the above relation

$$0\leq\Sigma\Sigma\,a_ia_j\,E(X_iX_j)\Rightarrow\Sigma\Sigma\,a_ia_j\,P(A_i\,A_j)\geq0.$$

10*. (i) $E(X^2)=(-1)^2.f(-1)+0^2.f(0)+1^2.f(1)=f(1)+f(-1)=[f(1)+f(0)+f(-1)]-f(0)=1-1/2=1/2$

(ii) $E(X)=(-1)f(-1)+0.f(0)+1.f(1)=f(1)-f(-1)=1/6\qquad[\because\ \mu=1/6]$

Using $f(0)=1/2$ gives $f(1)+f(-1)=1/2$, hence $f(1)=1/3$ and $f(-1)=1/6$.

11*. Let X denote the (random) number of face (picture) cards obtained in a Bridge-hand. then

$$P(X=x)=\binom{12}{x}\binom{40}{13-x}\Big/\binom{52}{13}$$

$$\therefore \quad \mathrm{E}(X) = \sum_{x=0}^{12} x\binom{12}{x}\binom{40}{13-x}\Big/\binom{52}{13} = 12\sum_{x=0}^{11}\binom{11}{x-1}\binom{40}{13-x}\Big/\binom{52}{13}$$

$$= 12\binom{51}{12}\Big/\binom{52}{13} = 3. \qquad \text{[By Hyp-geom Identity]}$$

12*. Here $\Omega = \{H, TH, TTH, \ldots\}$. Suppose the head occurs on kth trial, then

$$P(X = k) = q^{k-1}p, \qquad [X \sim \text{gem}(p)]\ k = 1, 2, 3, \ldots\ (p + q = 1).$$

where X is the number of trials (tosses) to get the first success (head).

Now $\mathrm{E}(X) = \sum_{k=1}^{\infty} k(q^{k-1}p) = p(1 + 2q + 3q^2 + \ldots) = p(1-q)^{-2} = p/p^2 = 1/p.$

For a fair coin, $P(H) = p = 1/2$, so $\mathrm{E}(X) = 2$.

13*. Let the variate X denote the number of failures occuring before the first success. Then (F = failure, S = success).

$$P(X = k) = P(FF \ldots FS) = P(F) \,.\, P(F) \ldots P(F)\, P(S) = q^k p,\ k = 0, 1, 2, \ldots$$

$$\therefore \quad \mathrm{E}(X) = \sum_{k=0}^{\infty} k.(pq^k) = pq\sum_{k=1}^{\infty} kq^{k-1}$$

$$= pq(1 + 2q + 3q^2 + 4q^3 + \ldots) = pq(1-q)^{-2} = pq/p^2 = q/p = (1/p) - 1.$$

14*. Let the variates X_i and X_j denote the face values of the two cards, $X_i \neq X_j$. The probability of the first extraction is $(1/n)$ and that of the second extraction is $1/(n-1)$. Hence

$$\mathrm{E}(X_i X_j) = \sum_i \sum_j \frac{x_i}{n} \cdot \frac{x_j}{n-1};\ i \neq j,\ i = 1, 2, 3, \ldots, n$$

$$\therefore \quad n(n-1)\,\mathrm{E}(X_i X_j) = \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{j\neq 1}^{n} x_j\right) = \left(\sum_{i=1}^{n} x_i\right)\left[\sum_{j=1}^{n} x_j - x_i\right] = \left(\sum_{i=1}^{n} x_i\right)^2 - \left(\sum_{i=1}^{n} x_i^2\right)$$

$$= [\tfrac{1}{2}n(n+1)]^2 - [n(n+1)(2n+1)/6] = n(n+1)(3n^2 - n - 2)/12 = n(n+1)(n-1)(3n+2)/12.$$

$$\therefore \quad \mathrm{E}(X_i X_j) = (n+1)(3n+2)/12.$$

15*. We use L.U.S. to get

$$\mathrm{E}(Y) = \mathrm{E}(X-1)^2 = 1.\frac{1}{3} + 0.\frac{1}{2} + 1.\frac{1}{24} + 4.\frac{1}{8} = \frac{21}{24} = \frac{7}{8}.$$

16*. Here $f_1(t) = 1/a$, $F_1(t) = t/a$. So, using survival function.

$$P(X > x) = P(X_1 > x)\, P(X_2 > x)\, P(X_3 > x) = [1 - P(X_1 \leq x)]^3 = [1 - (x/a)]^3$$

$$\therefore \quad 1 - F_X(x) = [1 - (x/a)]^3 \Rightarrow f_X = (3/a)[1 - (x/a)]^2, \ [\because F'(x) = f(x)]$$

So, $\quad \mathrm{E}[1 - (X/a)]^2 = (3/a)\int_0^a [1 - (x/a)]^4 dx = 3/5$

Sec. 5-42. Page 201

1*. Let H denote the height, in inches, of a student; B and G the event of a student being a boy or a girl. Then, by Multistage E-Rule, $\mathsf{E}(H) = P(B)\,\mathsf{E}(H \mid B) + P(G)\,\mathsf{E}(H \mid G)$.
Here $P(B) = 0.6$, $P(G) = 0.4$, $\mathsf{E}(H \mid B) = 65$, $G(H \mid G) = 60$.

$\therefore \quad \mathsf{E}(H) = (0.6)\,(65) + (0.4)\,(60) = 39 + 24 = 63''$.

2*. Designate the events : $A_i = \{X = i\}$, $1 \le i \le 5$, and thus $P(A_i) = 1/5$. The expected value of Y is given by Total-E rule :

$$\mathsf{E}(Y) = P(A_1)\,\mathsf{E}(Y \mid A_1) + P(A_2)\,\mathsf{E}(Y \mid A_2) + \ldots + P(A_5)\,\mathsf{E}(Y \mid A_5). \qquad \ldots(1)$$

Now $E(Y \mid A_1) = 1, \mathsf{E}(Y \mid A_2) = \left(\frac{1}{2}\right)1 + \frac{1}{2}(2) = 3/2, \mathsf{E}(Y \mid A_3) = \left(\frac{1}{3}\right)(1) + \left(\frac{1}{3}\right)(2) + \left(\frac{1}{3}\right)(3) = 2,$

$$E(Y \mid A_4) = \left(\tfrac{1}{4}\right)(1+2+3+4) = 5/2, P(Y \mid A_5) = (1/5)(1+2+3+4+5) = 3.$$

$$\therefore \qquad \mathsf{E}(Y) = (1/5)\,[1 + (3/2) + 2 + (5/2) + 3] = 2. \qquad \text{[by (1)]}$$

3*. Observe : $P(X = j \mid Y = k) = 1/k$, because the second card is selected from among the cards 1, 2, ..., k. The conditional distribution is symmetric about $X = \frac{1}{2}(k+1)$, which is the mid-point of the extreme values 1 and k.

This gives $\mathsf{E}(X \mid Y = k) = \frac{1}{2}(k+1)$. Also, $P(Y = k) = 1/n$ since Y can take up any one of the n values. Thus by total-E Rule :

$$\mathsf{E}(X) = \sum P(Y = y_i)\,\mathsf{E}(X \mid Y = y_i) = \sum_{k=1}^{n} P(Y = k)\,\mathsf{E}(X \mid Y = k) = \frac{1}{2n}\sum_{k=1}^{n}(k+1) = \frac{n+3}{4}.$$

Sec. 5-62. Page 210

1*. Let $X - \mu_X = X_0$, $Y - \mu_y = Y_0$, then

$$H = \mathsf{E}[(Y_0 - bX_0) + (\mu_Y - b\mu_X - a)]^2. \qquad \text{[Put } c = \mu_Y - b\mu_X - a\text{; use Bin. expansion]}$$

$$= \mathsf{E}\{(Y_0^2 + b^2X_0^2 - 2bX_0Y_0) + c^2 + 2c(Y_0 - bX_0)\}$$

$$= \mathsf{E}(Y_0^2) + b^2\mathsf{E}(X_0^2) - 2b\mathsf{E}(X_0Y_0) + c^2 \qquad \text{[by Lin E and by } \mathsf{E}(X_0) = 0 = \mathsf{E}(Y_0)\text{]}$$

$$= \sigma_Y^2 + b^2\sigma_X^2 - 2b\sigma_{XY} + c^2 \qquad \text{[by def. of Var- \& Cov.]} \qquad \ldots(1)$$

To minimize the value of H, use calculus rules :

$$\left(\frac{\partial H}{\partial a}\right) = 2c\left(\frac{\partial c}{\partial a}\right) = -2c;\ \left(\frac{\partial H}{\partial b}\right) = 2b\sigma_X^2 - 2\sigma_{XY},\ \left[\frac{\partial H}{\partial a} = 0 \Rightarrow c = 0\right] \qquad \ldots(2)$$

$$\frac{\partial^2 H}{\partial b^2} = 2\sigma_X^2 > 0; \left(\frac{\partial^2 H}{\partial a\,\partial b}\right) = 0, \left(\frac{\partial H}{\partial b}\right) = 0 \Rightarrow b = \sigma_{XY}/\sigma_X^2 = \rho\sigma_Y/\sigma_X \qquad \ldots(3)$$

$$\therefore \qquad \min H = \sigma_Y^2 + \rho^2\sigma_Y^2 - 2\rho^2\sigma_Y^2 = \sigma_Y^2(1-\rho^2) \qquad \text{[by (1) and (3)]}$$

2*. Let the joint p.m.f. of X and Y be as shown and a, b, are any positive numbers and $4p = 1$. Then, $P(X=0)=p$, $P(Y=0)=p$, $P(X=0, Y=0)=0$. Since $P(X=0, Y=0) \neq P(X=0), P(Y=0)$, the variates X and Y are not independent. Now

$X \downarrow$ \ $Y \rightarrow$	0	b	$2b$	$P(x)$
0	0	p	0	p
a	p	0	p	$2p$
$2a$	0	p	0	p
$P(y)$	p	$2p$	p	$4p$

$\mathrm{E}(X) = 0\,.\,p + a\,.\,2p + 2a\,.\,p = 4ap = a;$

$\mathrm{E}(Y) = 0\,.\,p + b\,.\,2p + 2bp = 4bp = b$

$\mathrm{E}(XY) = a(2b)p + 2ab)p = 4abp = ab$

$\Rightarrow \qquad \mathrm{E}(XY) = \mathrm{E}(X)\,\mathrm{E}(Y).$

3*. Let $X_i = 1$, if the head turns upon on the ith coin and $X_i = 0$, otherwise [indicator variate]. If X is the total number of heads on the n coins, then $X = X_1 + X_2 + \ldots + X_n$. Now

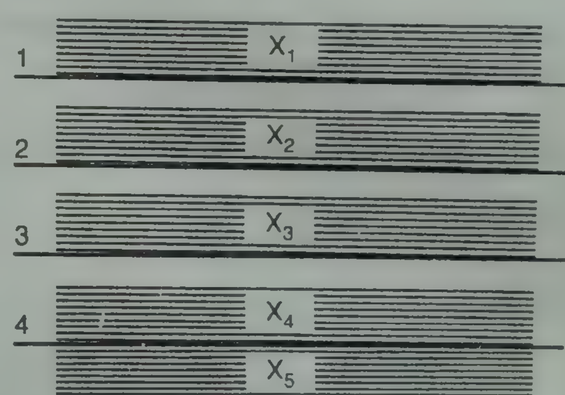
$$\mathrm{E}(X) = \mathrm{E}(X_1 + X_2 + \ldots + X_n) = \mathrm{E}(X_1) + \mathrm{E}(X_2) + \ldots + \mathrm{E}(X_n) = nE(X_1). \quad \text{[by Symmetry]}$$

If $P\{\text{head turns up}\} = p$, then $\mathrm{E}(X_1) = 1\,.\,p + 0\,.\,q = p$; hence $\mathrm{E}(X) = np$.

If the coins are unbiased, $p = 1/2$ and then $\mathrm{E}(X) = n/2$.

Comments. If four coins are tossed, then $\mathrm{E}(X) = 2$. If 5 coins are tossed, $\mathrm{E}(X) = 2.5$. This shows that the expected value of a discrete variate is not necessarily one of the possible values of the discrete variate, in which case, we would not "expect" to get the expected value. Hence the term *average value* or *mean value* is better than the term *Expected value.*

4*. Let X_1, X_2, X_3, X_4, X_5 be the number of non-ace cards as shown in the adjoining figure. Then, $X_1 + X_2 + X_3 + X_4 + X_5 = 48$.



$$\therefore \quad \mathrm{E}(X_1) + \mathrm{E}(X_2) + \ldots + \mathrm{E}(X_5) = 48 \qquad \ldots(1)$$

Now $\qquad \mathrm{E}(X_1) = \mathrm{E}(X_2) = \ldots = \mathrm{E}(X_5)$,

because any number of non-ace cards can be situated anywhere in the five spaces available. Hence, (1) gives, $5\mathrm{E}(X_1) = 48 \Rightarrow \mathrm{E}(X_1) = 48/5 = 9.6$.

5*. Let $X_i = 1$, if ball from urn i, $1 \le i \le 3$ is white; $X_i = 0$, otherwise. Now

$$\mathrm{E}(X_1) = 1\,.\,\tfrac{2}{5} + 0\,.\,\tfrac{3}{5} = \tfrac{2}{5};\ \mathrm{E}(X_2) = 1\,.\tfrac{11}{6} + 0\,.\,\tfrac{5}{11} = \tfrac{6}{11};\ \mathrm{E}(X_3) = 1\,.\,\tfrac{4}{6} + 0\,.\,\tfrac{2}{6} = \tfrac{2}{3}.$$

If $S = X_1 + X_2 + X_3$, is the number of white balls drawn, then

$$\mathrm{E}(S) = \mathrm{E}(X_1 + X_2 + X_3) = \mathrm{E}(X_1) + \mathrm{E}(X_2) + \mathrm{E}(X_3) = \tfrac{2}{5} + \tfrac{6}{11} + \tfrac{2}{3} = \tfrac{266}{165}.$$

6*. We use the indicators : Let $X_i = 1$, if ith ball drawn is white; $X_i = 0$, otherwise. If S is the number of white balls draw, then obviously $S = X_1 + X_2 + \ldots X_c$

$$\therefore \quad \mathrm{E}(S) = \mathrm{E}(X_1 + X_2 + \ldots + X_c) = \mathrm{E}(X_1) + \mathrm{E}(X_2) + \ldots + \mathrm{E}(X_c) = c\mathrm{E}(X_1) \qquad \text{[by symmetry]}$$

Now : $\ \mathrm{E}(X_i) = 1\,[a/(a=b)] + 0\,.[b/(a+b)] = a/(a+b).$

because, the probability that the ith ball removed is white, when nothing is known about the colours of the preceding balls is $a/(a+b)$. Hence, $\mathrm{E(S)} = ac/(a+b)$.

7*. Here $\Omega = \{1, 2, 3, 4, 5, 6\}$. Suppose integers X and Y are drawn in succession without replacement; then the p.m.f. of this drawing is $f(x, y) = (1/6)(1/5) = 1/30$. Now,

$E(|X - Y|) = 2E(X - Y), \ (X > Y)$. We find the set of such outcomes.

$$\{|X-Y|: X>Y\} = \{2-1, (3-1, 3-2), (4-1, 4-2, 4-3), (5-1, ..., 5-4), (6-1, ..., 6-5)\}$$

$$= \{1, (2, 1), (3, 2, 1), (4, 3, 2, 1), (5, 4, 3, 2, 1)\}$$

$\therefore \quad E(|X-Y|) = (2/30)[1+3+6+10+15] = 7/3.$ [See also Chap.1. p.14-15]

8*. Let X be the r.v. describing incme liable to tax so that X has the p.d.f. $f(x)$; this provides k] [by normallization]

$$1 = k\int_a^\infty x^{-\theta-1} dx \Rightarrow k = \theta a^\theta.$$

The static $t(x)$ describes the amount of tax, so that $t(x) = c(x - a)$. The average tax paid is $E[t(X)]$, hence

$$E[t(X)] = ck\int_a^\infty (x-a)\frac{dx}{x^{\theta+1}} = ck\int_a^\infty \left(\frac{1}{x^\theta} - \frac{a}{x^{\theta+1}}\right)dx = \frac{ck}{\theta(\theta-1)a^{\theta-1}} = \frac{ca}{\theta-1} \qquad ...(1)$$

Let the static $I(x)$ denote the increase in tax, then $I(x) = c(x - b),\ x \geq b$; $I(x) = 0,\ a \leq x < b$

$$\therefore \quad E[I(X)] = \int_b^\infty c(x-b)\frac{k\,dx}{x^{\theta+1}} = \frac{ck}{\theta(\theta-1)b^{\theta-1}} = \frac{ca^\theta}{(\theta-1)b^{\theta-1}}. \qquad \text{[by (1)]}$$

The percentage increase in average tax paid is 100 $E[I(X)]/E[t(x)]$ i.e. $100(a/b)^{\theta-1}$. This equal 20 provided $(b/a)^{\theta-1} = 5$.

9*. $$E(XY) = \sum_x\sum_y xy.(x+y^2)/42 = \sum_x\sum_y (x^2y + xy^3)/42$$

$$= (1/42)\sum_x (3x^2 + 27x)] = (1/42)(3.17 + 27.5) = 31/7$$

$$E(Y^2/X) = \sum_x\sum_y \frac{y^2}{x}\cdot\left(\frac{x+y^2}{42}\right) = \frac{1}{42}\sum_x\sum_y\left(y^2 + \frac{y^4}{x}\right) = \frac{1}{42}\sum_x\left(11 + \frac{83}{x}\right), \ [\Sigma y^2 = 11, \Sigma y^4 = 83]$$

$$= \frac{1}{42}\left[22 + \frac{415}{4}\right] = \frac{503}{168}. \ \left[\Sigma\frac{1}{x} = \frac{5}{4}\right]$$

10*. From definitions of expectation and c.d.f., we get

$$E[u(a - X)] = \int_{-\infty}^\infty u(a-x)f(x)\,dx = \int_{-\infty}^a f(x)\,dx = F_X(a),$$

Similarly, $E[u(b - Y)] = F_Y(b)$. Further,

$$E[u(a-X)u(b-Y)] = \int_{-\infty}^\infty\int_{-\infty}^\infty u(a-x)u(b-y)f(x,y)\,dx\,dy = \int_{-\infty}^a\int_{-\infty}^b f(x,y)\,dx\,dy = F(a,b)$$

If X and Y are independent, condition (1) is obviously true. If condition (1) holds, we get $F_{XY}(a, b) = F_X(a)\,F_Y(b) \Rightarrow X$ and Y are independent.

11*. As usual, $J_k(\omega) = 1$, if $\omega \in A_k$; $J_k(\omega) = 0$, if $\omega \notin A_k$. ...(1)

(i) Assume that events A_i, $1 \le i \le n$ are mutually independent. The definition (1) guarantees that $J_k(\omega)$ is a function of A_k. Thus the mutual independence of A_k's imply the mutual independence of J_k's.

(ii) Assume that the indicator variates $J_1, J_2, \dots, J_n$ are independent. Then,

$$E(J_1J_2 \dots J_n) = E(J_1)\,E(J_2) \dots E(J_n) \quad \text{...(2)}$$

Since $$E(J_1J_2 \dots J_n) = E(J_{A_1A_2\dots A_n}) = P(A_1A_2 \dots A_n) \quad \text{...(3)}$$

$$\therefore \quad P(A_1A_2 \dots A_n) = P(A_1)\,P(A_2) \dots P(A_n), \quad \text{[by (2) and (3)]}$$

We can similarly treat some or all complements of the events A_k's to obtain the other conditions for mutual independence, etc.

Sec. 5-75. Page 218

1*. Since X is partly discrete and partly continuous, we use

$$E(X-c)^k = k\int_0^\infty (x-c)^{k-1}[1-F(x)]\,dx - k\int_{-\infty}^{c}(x-c)^{k-1}F(x)\,dx \quad \text{...(1)}$$

Take $c = 0$, then (1) yields for the present case

$$E(X^k) = k\int_0^\infty x^{k-1}(pe^{-\lambda x})\,dx = pk[\Gamma(k)/\lambda^k]$$

$$\therefore \quad E(X) = p/\lambda, E(X^2) = 2p/\lambda^2, \text{ Var }(X) = p(2-p)/\lambda^2.$$

2*. $E(X) = 0p + 2p + 1(1-2p) = 1$; $E(X^2) = 0p + 4p + 1\cdot(1-2p) = 1 + 2p$.

$\text{Var}(X) = E(X^2) - E^2(X) = 2p$, $0 \le p \le 1/2$. Thus $\max[\text{Var}(X)] = 2 \max p = 2(1/2) = 1$.

3*. Let $X_i = +1$, if the ith step is towards the right; $X_i = -1$, if the ith step is towards the left. The, $E(X_i) = p(1) + q(-1) = p - q$, $\forall\, i$ ...(1)

Also $S = X_1 + X_2 + \dots + X_n$, represents the random distance from the origin after n steps. Now

$$E(S) = E(\Sigma X_i) = \Sigma E(X_i) = n(p-q), \quad \text{[by Lin E and using (1)]}$$

Also, $E(X_i^2) = p_1^2 + q(-1)^2 = p + q = 1$; $\text{Var }(X_i) = E(X_i^2) - E^2(X_i) = (p+q)^2 - (p-q)^2 = 4pq$

$$\therefore \quad \text{Var }(S) = \text{Var }(\Sigma X_i) = \Sigma \text{Var}(X_i) = 4npq$$

4*. Twice diffentiation of a G.P. : $1 + z + z^2 + z^3 + z^4 + \dots = (1-z)^{-1}$ gives

$1 + 2z + 3z^2 + 4z^3 + \dots = (1-z)^{-2}$, $1 + 3z + 6z^2 + 10z^3 + \dots = (1-z)^{-3}$. ...(1)

$$\therefore \quad E(X) = \sum_{k=1}^{\infty}(pq^{k-1})\,k = p\sum_{k=1}^{\infty} kq^{k-1} = p(1-q)^{-2} = 1/p, \quad \text{[by 1(a)]}$$

$$E[X(X-1)] = \sum_{k=1}^{\infty} pq^{k-1}.k(k-1) = 2pq(1 + 3q + 6q^2 + 10q^3 + \dots) = 2pq\,(1-q)^{-3} = 2p/p^2, \text{[by (1b)]}$$

i.e. $$E(X^2) - E(X) = 2q/p^2 \Rightarrow E(X^2) = (2q/p^2) + 1/p.$$

$$\text{Var }(X) = E(X^2) - E^2(X) = (q/p^2)$$

5*. The probability of getting a *four* is $p = 1/6$ and that of getting no *four* is $q = 5/6$. If the r.v. X denotes the number of throws to obtain a four, then the probability distribution is [S = Success, F = Failure].

Ω	:	S	FS	FFS	FFFS	...	FF ... FS
X	:	1	2	3	4	...	k
$f(x)$	:	p	qp	q^2p	q^3p	...	$q^{k-1}p$

The most probable number of throws corresponds to the maximum probability, which is $p = 1/6$. Hence the most probable number of throws is 1.

Further : $E(X) = 1/p$, $E[X(X-1)] = 2q/p^2$ [Ex. 5-27(a)]

With $p = 1/6$, $E(X) = 6$, $E(X^2) = 66$, $\text{Var}(X) = 66 - 36 = 30$.

6*. Let $AP = X$, then $PB = 2a - X$ and so the area of the rectangle $AP \cdot PB$ is $S = X(2a - X)$. Also, the point P is uniformly distributed over $AB = 2a$, so that $f(x) = 1/2a$, $0 < x < 2a$.

Now $\left\{S \geq \frac{1}{2}a^2\right\} = \{(X-a)^2 \leq \frac{1}{2}a^2\} = (a/\sqrt{2}) \leq X \leq a + (a\sqrt{2})$

$$\therefore \quad P\left(S \geq \tfrac{1}{2}a^2\right) = \int_{a-(a/\sqrt{2})}^{a+(a/\sqrt{2})} \frac{dx}{2a} = \frac{1}{\sqrt{2}}.$$

$$E(S) = E(2aX - X^2) = \int_0^{2a} (2ax - x^2)\frac{dx}{2a} = \frac{2a^2}{3}.$$

$$E(S^2) = E(X^4 - 4aX^3 + 4a^2X^2) = \int_0^{2a} (x^4 - 4ax^3 + 4a^2x^2)\frac{dx}{2} = 16a^5/15$$

$$\therefore \quad \text{Var}(S) = E(S^2) - E^2(S) = \frac{16a^5}{15} - \frac{4a^4}{9} = \frac{4a^4(12a-5)}{45}$$

7*. Here $f(\theta) = k\theta(\pi - \theta)$, $0 < \theta < \frac{1}{2}\pi$ where k is determined by normality :

$$1 = k\int_0^{\pi/2} (\pi\theta - \theta^2)\, d\theta = k\left[\pi\frac{\pi^2}{8} - \frac{\pi^3}{24}\right] = \frac{k\pi^3}{12}, \Rightarrow k = \frac{12}{\pi^3}$$

$$\therefore \quad f(\theta) = 12\theta(\pi - \theta)/\pi^3, 0 < \theta < \tfrac{1}{2}\pi.$$

The area of the said triangle is $S = \frac{1}{2}a^2 \sin\theta$; hence

$$E(S) = \frac{a^2}{2}\int_0^{\pi/2} \sin\theta \, \frac{12\theta\,(\pi-\theta)}{\pi^3}\, d\theta = \frac{6a^2}{\pi^3}[-(\pi\theta - \theta^2)\cos\theta + (\pi - 2\theta)\sin\theta - 2\cos\theta]_0^{\pi/2}$$
$$= 12a^2/\pi^3.$$

$$E(S^2) = \frac{3a^4}{2\pi^3}\int_0^{\pi/2} (\pi\theta - \theta^2)(1 - \cos 2\theta)\, d\theta \quad \{\because\ 2\sin^2\theta = 1 - \cos 2\theta\}$$

$$= \frac{3a^4}{2\pi^3}\left\{\frac{\pi}{2}\frac{\pi^2}{4} - \frac{\pi^3}{3\times 8}\right\} - \frac{3a^4}{2\pi^3}\int_0^{\pi/2} (\pi\theta - \theta^2)\cos 2\theta\, d\theta \qquad ...(1)$$

$$\text{Now, } \int_0^{\pi/2} (\pi\theta - \theta^2)\cos 2\theta\, d\theta = \left[(\pi\theta - \theta^2)\frac{\sin 2\theta}{2} + (\pi - 2\theta)\frac{\cos 2\theta}{4} + \frac{\sin 2\theta}{4}\right]_0^{\pi/2} = -\frac{\pi}{4}$$

Substituting into (1), we get

$$\mathsf{E}(S^2) = \frac{a^4}{8} + \frac{3a^4}{8\pi^2} = \frac{a^4}{8\pi^2}(3+\pi^2) - \frac{144a^4}{\pi^6} = a^4\left(\frac{\pi^6 + 3\pi^4 - 1152}{8\pi^6}\right)$$

$$\text{Var}(S) = \mathsf{E}(S^2) - \mathsf{E}^2(S) = \frac{a^4(3+\pi^2)}{8\pi^2}$$

8*. Associate with ith trial a variate X_i defined by

$X_i = 1$, if success occurs $P(X_i = 1) = p_i$; $X_i = 0$, if the failure occurs, $P(X_i = q_i) = q_i$.

$p_i + q_i = 1$. Then, $\mathsf{E}(X_i) = 1p_i + 0q_i = p_i, \mathsf{E}(X_i^2) = 1^2 . p_i + 0^2 . q_i = p_i; \text{Var}(X_i) = \mathsf{E}(X_i^2) - \mathsf{E}^2(X_i) = p_i - p_i^2$

Let S be the number of successes in n trials, then $S = X_1 + X_2 + \ldots + X_n$.

$\therefore \quad \mathsf{E}(S) = \mathsf{E}(X_1) + \mathsf{E}(X_2) + \ldots + \mathsf{E}(X_n) = \Sigma p_i, 1 \le i \le n$

$$\sigma^2 = \text{Var}(S) = \Sigma \text{Var}(X_i) = \Sigma(p_i - p_i^2) = \Sigma p_i - \Sigma p_i^2 \qquad \ldots(1)$$

When $p_i = p$, then $\Sigma p_i = np$, so that p is the mean of the quantities p_i. Let σ_0^2 be the variance of the quantities p_i, so that

$$\sigma_0^2 = [(\Sigma p_i^2)/n] - [(\Sigma p_i)/n]^2 \Rightarrow \Sigma p_i^2 = n(p^2 + \sigma_0^2) \qquad \ldots(2)$$

From (1) and (2), using $\Sigma p_i = np$, we get

$$\sigma^2 = np - n(p^2 + \sigma_0^2) = npq - n\sigma_0^2. \qquad \ldots(3)$$

Since $p_i = p$ (constant) yields Var $(X_i) = pq$, so in this case Var(S) $= \sigma_c^2 = npq$,

$\therefore \quad \sigma^2 = \sigma_c^2 - n\sigma_0^2 \;\Rightarrow\; \sigma_c^2 - \sigma^2 = n\sigma_0^2 > 0 \Rightarrow \sigma < \sigma_c.$

Chapter 6 : Moments and Quantiles

Sec. 6-13. Page 227

1*. Boundness $\Rightarrow$ $P\{|X| \le B\} = 1$. Assume that X is continuous with density $f(x)$. Then for any integer $n > 0$.

$\int_{-\infty}^{\infty} |x|^n f(x)\,dx \le \int_{-\infty}^{\infty} B^n f(x)\,dx = B^n \int_{-\infty}^{\infty} f(x)\,dx = B^n < \infty.$

Thus $\mathsf{E}(X^n)$ exists for all n. The proof for discrete X is no more different.

2*. Suppose X is continuous with p.d.f. $f(x)$. Since $\mu_k' = \mathsf{E}(X^k) < \infty$, the integral

$\int_{-\infty}^{\infty} x^k f(x)\,dx$ converges absolutely. Now

$$\mathsf{E}(|X|^k) = \int_{-\infty}^{\infty} |x|^k f(x)\,dx = \lim_{n\to\infty} \int_{|x|\le n} |x|^k f(x)\,dx \qquad \ldots(1)$$

Also $\mathsf{E}(|X|^k) = \int_{-\infty}^{\infty} |x|^k f(x)\,dx = \lim\left\{\int_{|x|>n} |x|^k f(x)\,dx + \int_{|x|\le n} |x|^k f(x)\,dx\right\} \qquad \ldots(2)$

Letting $n \to \infty$ in (2), and equating it to (1), we readily obtain

$$\lim_{n\to\infty} \int_{|x|>n} f(x)|x|^k\,dx = 0 \qquad \ldots(3)$$

But $\int_{|x|>n} |x|^k f(x)dx \ge n^k \int_{|x|>n} f(x)\,dx = n^k P(|X|>n)$ ...(4)

Letting $n \to \infty$ in (4), using (3) we get the result : $\lim n^k P\{\,|X|>n\} = 0$.

Converse of above result need not be true. Suppose a variate X is defined by

$$P(X=n) = k/n^2 \ln n, n=2,3,4,\ldots \qquad [\Sigma\,(k/n^2 \ln n)=1]$$

Then, $$P\{|X|>n\} = P\{X>n\} = \int_n^\infty \frac{kdx}{x^2 \ln x} \le \frac{k}{n \ln n} \qquad \ldots(5)$$

because $$\frac{1}{\ln x} < \frac{1}{\ln n} \Rightarrow \int_n^\infty \frac{dx}{x^2 \ln x} \le \frac{1}{\ln n}\int_n^\infty \frac{dx}{x^2} = \frac{1}{n \ln n}.$$

From (5), it follows that $nP\{\,|X|>n\} \le (k/\ln n) \to 0$ as $n \to \infty$.

But $E(X) = \Sigma nP(X=n) = \Sigma[k/n \ln n] = \infty$. When X is discrete, the proof is similar.

Sec. 6-25. Page 231

1*. $\mu_3 = \mu_3' - 3\mu_2'\mu_1' + 2\mu_2'^3 = \frac{1}{4}n(n+1)^2 - \frac{1}{4}(n+1)^2(2n+1) + \frac{1}{4}(n+1)^3 + 0.$

Thus $\mu_3 = 0$, $\Rightarrow \gamma_1 = 0$, so that skewness is zero.

2*. $\mu_n = E(X^n) = \int_{-\infty}^{\infty} x^n f(x)\,dx$. If n is odd, the integrand is an odd function of x, hence above integral vanishes. Thus

$$\mu_{2n} = 2\int_0^\infty x^{2n} f(x)\,dx = \frac{\lambda^\theta}{\theta\,\Gamma(\theta)}\int_0^\infty \theta\, z^{(2n+1)\theta-1} e^{-\lambda z}\,dz, \qquad [z = x^{1/\theta}]$$

$$= \Gamma[(2n+1)\theta]/\lambda^{2\pi\theta}\,\Gamma/(\theta).$$

$$\therefore \quad \beta_2 = \mu_4/\mu_2^2 = \Gamma(5\theta)\Gamma(\theta)/[\Gamma(3\theta)]^2.$$

3*. Here $E(X-2) = 1$, so that $\mu = E(X) = 3$; $\mu_1' = 1$. Now

$$\mu_2 = \mu_2' - \mu_1'^2 = 16 - 1 = 15;\ \mu_3 = \mu_3' - 3\mu_2'.\mu_1' + 2(\mu_1')^3 = -40 - 3\times 16 + 2 = -86.$$

Further : $16 = E(X-2)^2 = E(X^2) - 4E(X) + 4 = E(X^2) - 12 + 4 \Rightarrow E(X^2) = 24.$

$-40 = E(X-2)^3 = E(X^3 - 6X^2 + 12X - 8) = E(X^3) - 6\times 24 + 12\times 3 - 8 \Rightarrow E(X^3) = 76.$

Thus, first three simple moments are 1, 24, 76.

4*. With usual notation :

$$\mu_2 = \mu_2' - \mu_1'^2 = 4 - 1 = 3,\ \mu_3 = \mu'^3 - 3\mu_2'\mu_1' + 2(\mu_1')^3 = 10 - 3\times 4 + 2 = 0,$$

$$\mu_4 = \mu_4' - 4\mu_3'\mu_1' + 6\mu_2'\mu_1'^2 - 3\mu_1'^4 = 45 - 4\times 10 + 6\times 4 - 3 = 26.$$

Also, $E(X-4) = 1 \Rightarrow E(X) = \mu = 5$. $\beta_1 = \mu_3^2/\mu_2^3 = 0$, $\beta_2 = (\mu_4/\mu_2^2) = 26/9$.

Sec. 6-51. Page 237

1*. Let μ be the mean of the series that consists of $(2n + 1)$ terms, then

$$\mu = [a+(a+d)+(a+2d)+...+(a+2nd)]/(2n+1)$$

$$= [(2n+1)a+\tfrac{1}{2}(1+2n).2nd]/(2n+1) = a+nd$$

$$\delta(\mu) = \frac{\Sigma|(a+rd)-(a+nd)|}{(2n+1)} = \frac{|d|}{2n+1}\sum_{r=0}^{2n}|r-n| = \frac{2d}{2n+1}(1+2+3+...+n) = \frac{n(n+1)d}{2n+1}.$$

$$\sigma^2 = \frac{\Sigma[(a+rd)-(a+nd)]^2}{2n+1} = \frac{d^2}{2n+1}\sum_{r=0}^{2n}(r-n)^2 = \frac{2d^2}{2n+1}(1^2+2^2+3^2+...+n^2) = \frac{n(n+1)}{3}d^2$$

Let $D = \sigma^2 - [\delta(\mu)]^2 = \dfrac{n(n+1)d^2}{3} - \dfrac{n^2(n+1)^2d^2}{(2n+1)^2} = \dfrac{n(n+1)(n^2+n+1)d^2}{3(1n+1)^2}.$

Obviously, $D > 0 \Rightarrow \sigma^2 > \delta^2(\mu)$, i.e. $\sigma > \delta(\mu)$.

2*. Since M denotes M.a.D. about mean m of the frequency distribution, we get, by definition

$$M = \sum_i \frac{f_i}{N}|x_i - m| = \frac{1}{N}\Big[\sum_{x_i<m} f_i(m-x_i) + \sum_{x_i>m} f_i(x_i-m)\Big]$$

$$= N^{-1}\Big[\sum_{x_i<m} f_i(x_i-m) + \sum_{x_i>m} f_i(x_i-m)\Big] \qquad ...(1)$$

Also, $0 = E(X-m) = \Sigma\dfrac{f_i}{N}(x_i-m) = \dfrac{1}{N}\displaystyle\sum_{x_i<m} f_i(x_i-m) + \dfrac{1}{N}\sum_{x_i>m} f_i(x_i-m)$

To eliminate unwanted inequality $x_i > m_j$;

Subtract (2) from (1) to get

$$M = -(2/N)\sum_{x_i<m} f_i(x_i-m) = (2/N)\Big[m\sum_{x_i<m} f_i - \sum_{x_i<m} f_i x_i\Big]$$

3*. The d.f. for X is given by

$$F(x) = 0, \text{ if } x < 1; \quad F(x) = \sum_{k=1}^{[x]}\left(\tfrac{1}{2}\right)^k = 1-\left(\tfrac{1}{2}\right)^x, \text{ if } x \ge 1.$$

From $F(x) = 1-\left(\frac{1}{2}\right)^x$, we see that $F(1) = \frac{1}{2}$. Hence. the med $X = \frac{1}{2}(1+2) = \frac{3}{2}$.

The mean μ is given by : $\mu = E(X) = \Sigma x\left(\frac{1}{2}\right)^x, \quad x = 1, 2, ...$

Differentiating $(1-r)^{-1} = \Sigma r^x$, we get $(1-r)^{-2} = \Sigma xr^{x-1}; \quad x = 1, 2, ...$

$$\therefore \quad \mu = \tfrac{1}{2}\left(1-\tfrac{1}{2}\right)^{-2} = 2 \qquad \text{[using } r = \tfrac{1}{2} \text{ in above series]}$$

The mode of X is 1, because $f(1) = \frac{1}{2}$ is the maximum probability-ordinate.

Sec. 6-61. Page 241

1*. Let $\tau_0^2 = \tau_e^2$, then $\mathsf{E}(X - M_0)^2 = \mathsf{E}(X - M_e)^2$, which by Lin E, yields

$$2(M_0 - M_e)\,\mathsf{E}(X) = M_0^2 - M_e^2 \Rightarrow \mathsf{E}(X) = (M_0 + M_e)/2.$$

Converse. $2\mu = (M_0 + M_e) \Rightarrow 2\mu\,(M_0 - M_e) = M_0^2 - M_e^2$

$$\therefore \quad M_0^2 - 2\mu M_0 + \mathsf{E}(X^2) = M_e^2 - 2\mu M_e + \mathsf{E}(X^2) \Rightarrow \mathsf{E}(X - M_0)^2 = \mathsf{E}(X - M_e)^2 \Rightarrow \tau_0^2 = \tau_e^2$$

2*. Here : $\ln G = \mathsf{E}(\ln X) = \int_1^2 6(2-x)(x-1)\ln x\,dx = -6\int_1^2 (x^2 - 3x + 2)\ln x\,dx.$

Integrating by parts, we obtain

$$-\frac{1}{6}\ln G = \left[\left(\frac{1}{3}x^2 - \frac{3}{2}x^2 + 2x\right)\ln x\right]_1^2 - \int_1^2 \frac{1}{x}\left(\frac{1}{3}x^2 - \frac{3}{2}x^2 + 2x\right)dx$$

$$= \frac{2}{3}\ln 2 - \left[\frac{1}{9}x^3 - \frac{3}{4}x^2 + 2x\right]_1^2 = \frac{2}{3}\ln 2 - \frac{19}{36} = \frac{24\ln 2 - 19}{36}$$

$$\therefore \quad 6\ln G + 6\ln 2^4 = 19 \Rightarrow 6\ln(16G) = 19 \Rightarrow G \neq (\tfrac{1}{16})\exp(\tfrac{19}{6}).$$

3*. Let $Y = X - m$, so that Y is small to the extent that Y^3 is negligible. Now $X = Y + m$ provides, using Binomial Expansion

$$X^n = (m)^n[1 + Y/m]^n = (m)^n[1 + (nY/m) + n(n-1)Y^2/2m^2 + \ldots]$$

$$\therefore \quad \mathsf{E}(X^n) = (m)^n\,\mathsf{E}\{1 + (nY/m) + [n(n-1)/2m^2]Y^2 + \ldots\} \qquad \ldots(1)$$

Using the linearity of E, setting $\mathsf{E}(Y) = 0$, $\mathsf{E}(Y^2) = \sigma^2$, (1) reads

$$\mathsf{E}(X^n) = (m)^n\,\{1 + n(n-1)\sigma^2/2m^2\}$$

Putting $n = \frac{1}{2}$ and $-\frac{1}{2}$ in (2) we get

$$\mathsf{E}(X^{1/2}) = \sqrt{m}\left[1 - \frac{\sigma^2}{8m^2}\right], \quad \mathsf{E}\left(\frac{1}{\sqrt{X}}\right) = \frac{1}{\sqrt{m}}\left[1 + \frac{3\sigma^2}{8m^2}\right].$$

4*. We can determine C by normality $\Sigma f(x) = 1$. Now

$$F_X(x) = \int_{-\infty}^{x}\frac{C\,dt}{1+t^2} = C[\tan^{-1}t]_{-\infty}^{x} = C\left[\frac{\pi}{2} + \tan^{-1}x\right].$$

Since $F(\infty) = 1$, the above yields : $1 = C\pi \Rightarrow C = 1/\pi$. Thus

$$F(x) = (\tfrac{1}{2}) + (1/\pi)\tan^{-1}x, \quad -\infty < x < \infty. \qquad \ldots(1)$$

(a) $P(X \geq 0) = 1 - P(X \leq 0) = 1 - F(0) = 1 - \frac{1}{2} = \frac{1}{2}$ [by (1)]

(b) Using $F(Q_r) = r/4$ in (1), we get

$$\tan^{-1}(Q_r) = \pi(r-2)/4 \Rightarrow Q_r = -\cot(\pi r/4)$$

$$\therefore \quad Q_1 = -1,\; Q_2 = 0,\; Q_3 = 1,\; Q_3 - Q_1 = 2 \quad \text{(quartile range)}$$

(c) $$E(|X|) = C\int_{-\infty}^{\infty}\frac{|x|\,dx}{1+x^2} = 2C\int_0^{\infty}\frac{x\,dx}{1+x^2} \to \infty.$$

It follows that $E(X)$ does not exist. Consequently, variance also does not exist.

(d) Mode can be determined by 2nd derivative test. However, in this case

$$\text{mode } f(x) = \max f(x) = \min (f(x))^{-1} = \min (1 + x^2)/C.$$

Since $\min (x^2) = 0$, so $x = 0$ yields the modal value which is $1/\pi$.

5*. (i) $E(X^r) = \int_0^2 x^r f(x)\,dx = \int_0^1 x^{r+3}\,dx + 3\int_1^2 x^r(2-x)^3\,dx = (r+4)^{-1} + 3\int_0^1 t^3(2-t)^r\,dt.\ [2-x=t]$

$$\therefore \quad E(X) = \frac{1}{5} + 3\int_0^1 (2t^3 - t^4)\,dt = \frac{1}{5} + \frac{9}{10} = \frac{11}{10}$$

$$E(X^2) = \frac{1}{6} + 3\int_0^1 (t^5 - 4t^4 + 4t^3)\,dt = \frac{1}{6} + \frac{11}{10} = \frac{38}{30}$$

$$\therefore \quad \sigma^2 = \mu_2' - (\mu_1')^2 = \frac{38}{30} - \frac{121}{100} = \frac{17}{300}. \quad \Rightarrow \sigma = \left(\frac{17}{300}\right)^{1/2}$$

$$M = E\left\{\left|X - \frac{1}{20}\right|\right\} = \int_0^1 \left|x - \frac{11}{10}\right| x^3\,dx + 3\int_1^2 \left|x - \frac{11}{10}\right| (2-x)^3\,dx$$

$$= \int_0^1 \left(\frac{11}{10} - x\right) x^3\,dx + 3\int_0^1 \left|\frac{9}{10} - t\right| t^3\,dt, \quad [t = 2-x]$$

$$= \frac{3}{40} + 3\int_0^{0.9} t^3\left(\frac{9}{10} - t\right)dt + 3\int_{0.9}^1 t^3\left(t - \frac{9}{10}\right)dt$$

$$= \frac{3}{40} + 3\left[-\frac{1}{40} + \frac{1}{20}a\right] + 3\frac{a}{20} = \frac{3}{10}\cdot\left(\frac{9}{10}\right)^5 \quad [a = (9/10)^5]$$

(ii) This is similar to (i).

6*. The constant a shall be determined by Normality. Now

$$E(X^r) = \int_0^2 f(x)\,x^r\,dx = a\int_0^1 x^{x+2}\,dx + a\int_1^2 x^r(2-x)^2\,dx$$

$$= \frac{a}{r+3}[x^{r+3}]_0^1 + a\left[\frac{4x^{r+1}}{r+1} + \frac{x^{x+3}}{r+3} - 4\frac{x^{r+2}}{r+2}\right]_1^2 = 4a\left[\frac{2^{r+1}}{r+3} + \frac{2^{r+1}}{r+1} - \frac{2^{r+2}}{r+2}\right] \qquad \ldots(1)$$

To find a, put $r = 0$, use $E(1) = 1$ to get : $1 = 4a\,[(\frac{2}{3}) + 1 - (\frac{3}{2}) = 4a/6 \Rightarrow a = \frac{3}{2}$.

We now put $r = 1, 2, 3, 4, \ldots$ to obtain initial moments from (1).

$$E(X) = 6\left[\frac{4}{4} + \frac{3}{2} - \frac{7}{3}\right] = 1, \qquad E(X^2) = 6\left[\frac{8}{5} + \frac{7}{3} - \frac{15}{4}\right] = \frac{11}{10}.$$

$$E(X^3) = 6\left[\frac{8}{3} + \frac{15}{4} - \frac{31}{5}\right] = \frac{13}{10}, \qquad E(X^4) = 6\left[\frac{32}{7} + \frac{31}{5} - \frac{63}{6}\right] = \frac{57}{35}.$$

We may obtain first four central moments using $E(X^r)$, $r = 1, 2$ and $3, 4$. However, since $E(X) = 1$, we can use

$$\mathrm{E}(X-1)^r = a\int_0^1 (x-1)^r x^2\,dx + a\int_1^2 (x-1)^r (2-x)^2\,dx \qquad [\text{Put } x-1=y]$$

$$= a\int_{-1}^0 y^r(y^2+2y+1)\,dy + a\int_0^1 y^r(y^2-2y+1)\,dy.$$

i.e. $$\mu_r = a\int_0^1 y^r(y^2-2y+1)[1+(-1)^r]\,dy \quad [\text{Change } y \to -y \text{ in first integral}]$$

When r is an odd integer, we get $\mu_{2n+1}=0$. Putting $r=2n$ yields

$$\mu_{2n} = 2a\int_0^1 y^{2n}(y^2-2y+1)\,dy = 3\left\{\frac{1}{2n+3}-\frac{2}{2n+2}+\frac{1}{2n+1}\right\}$$

$$\therefore \quad \mu_2 = \frac{1}{10}=\sigma^2;\ \frac{\mu_4}{4}=\frac{1}{35} \Rightarrow S_k=\frac{\mu_3}{\sigma^3}=0,\ \text{Excess}=\frac{\mu_4}{\sigma^4}-3=\frac{100}{35}-3=-\frac{1}{7}.$$

7*. Assuming that given law is a p.d.f. we find that

$$\mathrm{E}(X-a)^r = y_0\int_a^\infty e^{-b(x-a)}(x-a)^r = y_0\int_0^\infty e^{-bz}\,z^{(r+1)-1}dz = y_0\cdot\frac{\Gamma(r+1)}{b^{r+1}} = y_0\frac{r!}{b^{r+1}}\ [x-a=z] \quad \ldots(1)$$

For f to be a density function, $\sum f(x)=1$ (Normality). Hence putting $r=0$, the above provides $1=y_0/b \Rightarrow y_0=b$ and so

$$\mu'_r = \mathrm{E}(X-a)^r = r!/b^r. \qquad \ldots(1)$$

When $r=1$, Eq. (1) gives $\mathrm{E}(X)-a=1/b$ or $m=a+b^{-1}$.

$$\mu_2 = \mu'_2-\mu_1'^2 = (2/b^2)-(1/b)^2 = 1/b^2 \Rightarrow \sigma=1/b.$$

Combining : $$y_0 = b = 1/\sigma,\ a = m-\sigma.$$

$$\mu_3 = \mu'_3-3\mu'_2\mu'_1+2\mu_1'^3 = 2/b^3 = 2\sigma^3.$$

$$\mu_4 = \mu'_4-4\mu'_3\mu'_1+6\mu'_2\mu_1'^2-3\mu_1'^4 = 9/b^4 = 9\sigma^4.$$

$$\therefore \quad \beta_1 = \frac{\mu_3^2}{\mu_2^3}=\frac{4\sigma^6}{\sigma^6}=4;\quad \beta_2=\frac{\mu_4}{\mu_2^2}=\frac{9\sigma^4}{\sigma^4}=9.$$

8*. We shall determine y_0 by normality $\sum f(x)=1$. Now

$$\mathrm{E}(X^r) \equiv \mu_r = y_0\int_0^1 x^r(x-x^2)\,dx\ \frac{y_0}{(r+2)(r+3)}.$$

Putting $r=0$, $\mu'_0\equiv 1=y_0/6$, we get $y_0=6$. Now let r 1, 2, to get

$$\mu'_1 = \mathrm{E}(X)=\tfrac{1}{2}, \mu'_2=\tfrac{3}{10},\ \text{so } \sigma^2=\mu'_2-\mu_1'^2=\tfrac{3}{20}.$$

$$(1/H) = \mathrm{E}(X^{-1})=3 \Rightarrow H=\tfrac{1}{3}.$$

$$\ln G = y_0\int_0^1(x-x^2)\ln dx$$

$$\therefore \quad \frac{1}{6}\ln G = \left[\left(\frac{x^2}{2}-\frac{x^3}{3}\right)\ln x\right]_0^1 - \int_0^1\left(\frac{x}{2}-\frac{x^2}{3}\right)dx = -\frac{5}{36}$$

$$G = \exp(-5/6).$$

For Mode : $f'(x) = 6(1-2x)$, $f''(6) = -12$ so $f'(x) = 0 \Rightarrow x = \frac{1}{2}$.

$$f''(\tfrac{1}{2}) = -12 < 0. \text{ Hence } x = \tfrac{1}{2} \text{ is the Modal value.}$$

Let the medium be m, then by its defnition

$$\tfrac{1}{2} = \int_0^m 6(x-x^2)\,dx = [3x^2 - 2x^3]_0^m \Rightarrow 4m^3 - 6m^2 + 1 = 0.$$

So, $(2m-1)(2m^2-2m-1) = 0$. This gives $m = \frac{1}{2} \in [0, 1]$. Hence $m = Q_2 = \frac{1}{2}$.

Since Mode = Median = Mean = $\frac{1}{2}$, the distribution is symmetrical.

To. determine M.D., we see that

$$\text{M.D.} = \mathrm{E}\left(\left|X-\tfrac{1}{2}\right|\right) = 6\int_0^1 |x-\tfrac{1}{2}|\, 6x(1-x)\,dx, \quad (\textit{Put } x-\tfrac{1}{2} = z)$$

$$= 6\int_{-1/2}^{1/2} |z|\left(\frac{1}{4}-z^2\right)dz = 12\int_0^{1/2} z\left(\frac{1}{4}-z^2\right)dz = \frac{12}{64} = \frac{3}{16}.$$

Note. **We can find central moments directly :**

$$\mu_r = \mathrm{E}\left(X-\frac{1}{2}\right)^r = 6\int_0^1\left(x-\frac{1}{2}\right)^r x(1-x)\,dx = 6\int_{-1/2}^{1/2} z^r\left(\frac{1}{4}-z^2\right)dz.$$

If $r = 2n+1$, the integrand is an odd function and so integral vanishes. Hence $\mu_{2n+1} = 0$.
If $r = 2n$, the integrand is even, and we get

$$\mu_{2n} = 12\int_0^{1/2} z^{2n}\left(\frac{1}{4}-z^2\right)dz = \frac{3}{2^{2n}(2n+1)(2n+3)}.$$

9*. $\mathrm{E}(X^r) = \int_0^\infty x^r \cdot \dfrac{xe^{-x^2/2a^2}}{a^2}\,dx = (2a^2)^{r/2}\int_0^\infty z^{(r/2+1)-1}e^{-z}\,dz \qquad \left[\dfrac{x^2}{2a^2} = z\right]$

So $\mu_r' = \mathrm{E}(X^r) = (2a^2)^{r/2}\,.\,\Gamma(\tfrac{1}{2}r+1)$.

$\therefore \mu_1' = a\sqrt{2}\,\Gamma(3/2) = a(\pi/2)^{1/2}$; $\mu_2' = 2a^2\,\Gamma(2) = 2a^2 . \sigma^2 = 2a^2 - (\pi a^2/2)$, $\Rightarrow \sigma = a\sqrt{(4-\pi)/2}$. ...(1)

If Q_r is the rth quartile, then

$$\frac{r}{4} = \int_0^{Q_r}\frac{te^{-t^2/2a^2}}{a^2}\,dt = \left[e^{-t^2/2a^2}\right]_{Q_r}^0 = 1 - e^{-(Q_r)^2/2a^2}$$

This gives, $(Q_r)^2 = -2a^2 \ln[1-(r/4)]$, whence $Q_1^2 = 2a^2 \ln(\frac{4}{3})$, $Q_3^2 = 2a^2 \ln 4$.

Using (1), $\dfrac{Q_3 - Q_1}{\sigma} = \dfrac{2\{(\ln 4)^{1/2} - [\ln(4/3)]^{1/2}\}}{\sqrt{4-\pi}}$ [Indep. of a]

10*. The constant a is determined by Normality $\Sigma f(x) = 1$. Now

$$\mathrm{E}(X^r) = a\int_{-3}^{-1}(3+x)^2 x^r\,dx + a\int_{-1}^{1}(6-2x^2)\,x^r\,dx + a\int_1^3 (3-x)^2 x^r\,dx.$$

Putting $x = -t$ in the first integral, transposing a to L.H.S., we get

$$a^{-1}\,\mathsf{E}(X^r) = \int_1^3 (3-t)^2\,(-t)^r\,dt + \int_1^3 (3-x)^2\,(3-x^2)\,x^r\,dx + \int_{-1}^1 (6-2x^2)\,x^r\,dx$$

$$= \int_1^3 (3-x)^2\,[1+(-1)^r]\,x^r\,dx + \int_{-1}^1 (6-2x^2)\,x^r\,dx. \text{ [Dummy } t \to \text{dummy } x] \quad ...(1)$$

When $r = 0$, this yields the value of a

$$a^{-1} = 2\int_1^3 (3-x)^2\,dx + 2\int_0^1 (6-2x^2)\,dx = 2\left[\frac{8}{3}+\frac{16}{3}\right] = 16.$$

Thus $a = \frac{1}{16}$. Further for odd $r = 2n - 1$, both integrals in (1) vanish, hence $\mathsf{E}(X^{2n-1})$ $= 0$. Thus, all odd-order simple moments vanish. In particular, $\mathsf{E}(X) = \mu = 0$, so that $\sigma^2 = \text{Var}(X) = \mathsf{E}(X^2)$. So, we put $r = 2$ in (1) to get

$$16\sigma^2 = 2\int_1^3 x^2(3-x)^2\,dx + 2\int_0^1 (6x^2-2x^4)\,dx$$

i.e.
$$\sigma^2 = \frac{1}{8}\int_1^3 (x^4-6x^3+9x^2)\,dx + \frac{1}{8}\int_1^0 (6x^2-2x^4)\,dx = \frac{1}{8}\left(\frac{64}{10}+\frac{8}{5}\right) = 1.$$

$$\text{M.D.} = \mathsf{E}(|X|) = \int_{-3}^3 |x|\,f(x)\,dx$$

$$= a\int_{-3}^{-1} |x|(3+x)^2\,dx + a\int_{-1}^1 |x|(6-2x^2)\,dx + a\int_1^3 |x|(3-x)^2\,dx$$

$$= 2a\int_1^3 x(x^2-6x+9)\,dx + 2a\int_0^1 (6x-2x^3)\,dx = 2a(4+\tfrac{5}{2}) = \tfrac{13}{16}.$$

11*. Here $f_X(x) = F_X'(x) = (a/c).(c/x)^{a+1}$, if $x > c$, $f(x) = 0$ if $x \le c$.

$$\mathsf{E}(X^r) = \frac{a}{c}(c^{a+1})\int_c^\infty x^{r-a-1}dx \quad \frac{ac^r}{a-r} \quad \text{provided } a > r.$$

$$\therefore \quad \mathsf{E}(X) = \frac{ac}{a-1},\ \mathsf{E}(X^2) = \frac{ac^2}{a-2},\ \text{Var}(X) = \frac{ac^2}{(a-1)^2(a-2)}. \quad [\mu_2 = \mu_2' - \mu^2]$$

The median m is the root of the equation $F(m) = \frac{1}{2} = 1-(c/x)^a \Rightarrow (c/x) = (\frac{1}{2})^{1/a}$

Thus $x = c\ (2)^{1/2}$ is the median of the distribution.

Obviously, for $a \ge 2$, we have $2^{1/a} < a/(a-2) \Rightarrow m < \mu$. Further, since $f'(x) < 0$ for $x > c$, the density function is monotone decreasing for $x > c$. As such, mode $m_0 = c$.

The quantile t_p of order $p = 0.75$ is the solution of the equation $P(X < t_p) = p = 0.75$.

i.e. $1-(c/t_p)^a = p \Rightarrow (c/t_p) = (1-p)^{1/a} = (0.25)^{-1/a} \Rightarrow t_{0.75} = c(0.25)^{-1/a}$

With stated values : $t_{0.75} = 100/(0.25)^{1/3} = 156.2.$ (Using ln tables)

Sec. 6-75. Page 250

1*. $\text{Var}(X+Y) = \sigma_X^2 + \sigma_Y^2 + 2\sigma_{XY} = 2+2-6 = -2$

Since Var $(X + Y) \not< 0$, the given data is not possible.

2*. Here $E(XY) \neq E(X)\,E(Y)$, so using Lin E, we get

$$Q = [E(XY) + E(Y)] - [E(X) + 1]E(Y) = E(XY) - E(X)E(Y) = \text{Cov}(X, Y)$$

Note. If X and Y are independent, then Cov $(X, Y) = 0$.

3*. For indicator variates, $E(X) = P(X = 1)$, $E(XY) = P\{X = 1, Y = 1\}$

$\therefore$ $\quad$ $\text{Cov}\,(X, Y) = E(XY) - E(X)\,E(Y) = P\{X = 1, Y = 1\} - P\{X = 1\}\,P\{Y = 1\}$

Thus $\quad \sigma_{XY} = 0$ iff $P(X = 1, Y = 1) = P(X = 1)\,P(Y = 1)$.

This shows that X and Y are independent. [Other possibilities $X = 0$, $Y = 0$, etc. trivially hold]

Note. Cov $(I_A, I_B) = 0$, iff A and B are independent.

4*. Let $X_k = 1$, if card k falls in cell k and $X_k = 0$, otherwise (X_k : indicator r.v.). The probability of any card falling in cell k is $1/n$. Now

$E(X_k) = 1\,.\,P(X_k = 1) + 0\,.\,P(X_k = 0) = 1/n\,;\, E(X_k^2) = 1^2\,.\,P(X_k = 1) + 0^2\,.\,P(X_k = 0) = 1/n$

$\text{Var}\,(X_k) = E(X_k^2) - E^2(X_k) = (1/n) - (1/n^2) = (n-1)/n^2$

$E(X_k X_j) = 1/n(n-1)$, $\text{Cov}(X_j X_k) = [1/n\,(n-1)] - (1/n^2) = 1/n^2\,(n-1)$.

Notice that $X_k\, X_j = 1$ iff card k falls in cell k and j falls in cell j, the former with probability $1/n$ and the later with probability $1/(n-1)$; and in all other cases $X_k\, X_j = 0$. If S denotes the total number of matches, then $S = X_1 + X_2 + \ldots + X_n$.

$$\therefore \qquad E(S) = E(X_1)E(X_2) + \ldots + E(X_n) = nE(X_i) = 1$$

$$\text{Var}\,(S) = \text{Var}\,(X_1 + \ldots + X_n) = \sum \text{Var}\,(X_i) + 2\sum \text{Cov}\,(X_i, X_j), \qquad j > i$$

$$= \frac{n(n-1)}{n^2} + 2\,.\binom{n}{2}\frac{1}{n^2\,(n-1)} = \frac{n-1}{n} + \frac{1}{n} = 1.$$

5*. Let $X_i = 1$, if ith ball drawn is green, $X_i = 0$, otherwise ; $i = 1, 2, \ldots, k$. Then, $S_k = X_1 + X_2 + \ldots + X_k$, is the number of green balls drawn. Now

$$P(X_i = 1) = \frac{m}{m+n},\; P(X_i = 0) = \frac{n}{m+n},\; E(X_i) = P(X_i = 1) = \frac{m}{m+n} = E(X_i^2).$$

$\therefore \quad \text{Var}\,(X_i) = [m/(m+n)] - [m/(m+n)]^2 = mn/(m+n)^2$.

To compute Cov (X_i, X_j), $i \neq j$ we observe that $X_i\, X_j = 1$, if both ith and jth balls drawn are green ; otherwise $X_i\, X_j = 0$. Hence

$E(X_i\, X_j) = P(X_i = 1, X_j = 1) = P(X_i = 1)\,P(X_j = 1 \mid X_i = 1) = [m/(m+n)]\,[(m-1/(m+n-1)]$.

$$\text{Cov}\,(X_i, X_j) = E(X_i\, X_j) - E(X_i)E(X_j)$$

$$= \frac{m(m-1)}{(m+n)\,(m+n-1)} - \left(\frac{m}{m+n}\right)^2 = \frac{-mn}{(m+n)^2\,(m+n-1)}.$$

$$E(S_k) = k\,E(X_i) = mk/(m+n).$$

$$\text{Var}\,(S_k) = \Sigma\,\text{Var}(X_j) + 2\Sigma\,\text{Cov}\,(X_i, X_j) \;\; (i \neq j\,;\, i, j = 1, 2, \ldots, k)$$

$$= \frac{kmn}{(m+n)^2} - 2\binom{k}{2}\frac{mn}{(m+n)^2\,(m+n-1)} = \frac{mnk\,(m+n-k)}{(m+n)^2\,(m+n-1)}.$$

6*. Let the numbers appearing on the n successively drawn tickets be $X_1, X_2, ..., X_n$. These are jointly distributed. Now

$$\text{Var}\,(S_n) = \text{Var}(X_1 + ... + X_n) = \sum_{j=1}^{n} \text{Var}(X_i) + 2\sum_{j \neq k} \text{Cov}\,(X_j, X_k). \qquad ...(1)$$

Observe that $\quad \bar{v} = (\Sigma v_j)/N; \quad \sigma^2 = \Sigma(v_j - \bar{v})^2/N = \Sigma(v_j^2/N) - (\bar{v})^2.$

Now Var $(X_j) = \sigma_2$, $1 \le j \le n = 1, 2, ..., n$. Further, each of the ${}^nC_2 = n(n-1)/2$ pairs of X_s has the same joint distribution and hence the same covariance (say) θ. Thus, (1) provides

$$\text{Var}\,(S_n) = n\sigma^2 + n(n-1)\,\theta, \quad n \le N.$$

To determine θ, employ a trick : take $n = N$, so that $S_N = \Sigma X_j = \Sigma v_j$ = const. This means Var $(S_N) = 0$ and from (2), we get $\theta = -\sigma^2/(N-1)$ and thus (2) begets

$$\text{Var}\,(S_n) = n\sigma^2 - [n(n-1)\,\sigma^2/(N-1)] = [n(N-n)/(N-1)]\,\sigma^2.$$

$$\text{Var}(\bar{X}_n) = \text{Var}\,(S_n/n) = (1/n^2)\,\text{Var}(S_n) = [(N-n)/(N-1)]\,(\sigma^2/n).$$

The factor $(N-n)/(N-1)$ is the correction factor for sampling without replacement [Var $(S_n/n) = \sigma^2/n$ for the case of sampling with replacement]

$$\text{Corr}\,(X_j, X_k) = \text{Cov}(X_j, S_k)/\sqrt{\text{Var}(X_i)\,\text{Var}(X_k)} = -1/(N-1).$$

Note. If $v_j = j$ then $\sigma^2 = (N^2-1)/12$, since $\bar{v} = (N+1)/2$.

APPENDIX D. Sec. D-5. Page 259

1*. For the numeral $x_k = k$, the weight $w_k = k$. By definition, the weighted mean is

$$\bar{x}_w = \frac{\Sigma w_i\,x_i}{\Sigma x_i} = \frac{\Sigma k^2}{\Sigma k} = \frac{n(n+1)\,(2n+1)/6}{n(n+1)/2} = \frac{(2n+1)}{3}.$$

2*.

$$E(X) = \Sigma(x_i/n) = (1+2+...+n)/n = (n+2)/2$$

$$E(X^2) = \Sigma x_i^2/n = (1^2+2^2+...+n^2)/n = n(n+1)\,(2n+1)/6n$$

$$\text{Var}\,(X) = E(X^2) - E^2(X) = \frac{(n+1)\,(2n+1)}{6} - \frac{(n+1)^2}{4} = \frac{n^2-1}{12}.$$

Remark. For a Die distribution, $n = 6$ so that $\mu = 7/2$ and $\sigma^2 = 35/12$.

3*. We are given that, if k is some constant, then

$$p_r = P(X = r) = k\binom{n}{r}, \quad r = 0, 1, 2, ..., n$$

$$\therefore \quad 1 = \Sigma\,p_r\,k\Sigma\binom{n}{r} = k(1+1)^n = k.2^n \;\Rightarrow\; k = (1/2)^n.$$

$\mathrm{E}(X) = \Sigma r\, p_r = k\,\Sigma\binom{n}{r} r = nk \sum_1^n \binom{n-1}{r-1} = nk\,(1+1)^{n-1} = n\,.\,2^{n-1}/2^n = n/2.$ $\left[\binom{n}{r} = \frac{n}{r}\binom{n-1}{r-1}\right]$

$\mathrm{E}(X^2) = \Sigma r^2 p_r = k\,\Sigma[r(r-1)+r]\binom{n}{r} = k[n(n-1)^{2n-2} + n2^{n-1}] = n(n+1)/4.$

$$\mathrm{Var}\,(X) = \mathrm{E}(X^2) - \mathrm{E}^2(X) = n/4.$$

4*. $(\Sigma x_i)_c$ = Corrected $\Sigma x_i = \Sigma x_i - x_1 + y_1 = T - x_1 + y_1$ ...(1)

$(\Sigma x_i^2)_c$ = Corrected $\Sigma (x_i)^2 = \Sigma x_i^2 - x_1^2 + y_1^2$ ...(2)

Originally : $\sigma^2 = (\Sigma x_i^2/n) - [\Sigma x/n]^2$, which provides $(\Sigma x_i^2)/n = \sigma^2 + (T^2/n^2)$. ...(3)

Let σ_c^2 be the corrected variance, then using (1) and (2)

σ_c^2 = Corrected $(\Sigma x_i^2/n)$ – [Corrected $(\Sigma x_i/n)]^2 = [((\Sigma x_i^2 - x_1^2 + y_1^2)/n] - [(T - x_1 + y_1)/n]^2$,

$$= \sigma^2 + \frac{T^2}{n^2} + \frac{y_1^2 - x_1^2}{n} - \left[\frac{T}{n} + \frac{y_1 - x_1}{n}\right]^2$$

$$\therefore \quad \sigma_c^2 - \sigma^2 = \frac{y_1^2 - x_1^2}{n} - \left[\left(\frac{y_1 - x_1}{n}\right)^2 + \frac{2T(y_1 - x_1)}{n^2}\right] = \frac{y_1 - x_1}{n}\left[(y_1 + x_1) - \frac{y_1 - x_1 + 2T}{n}\right].$$

5*. Let the values of variate X be $x_1, x_2, ..., x_n$ and let

$$Q_p = [(x_1^p + x_2^p + ... + x_n^p)/n]^{1/p} \quad ...(1)$$

(a) Choose $p = 1$ in (1) to get

$$Q_1 = (x_1, x_2, ..., x_n/n) = \bar{x} \text{ (the arithmetic mean).}$$

(b) Choose $p = -1$ in (1) to get

$$Q_{-1} = \left[\left(\frac{1}{x_1} + \frac{1}{x_2} + ... + \frac{1}{x_n}\right)\frac{1}{n}\right]^{-1} = \frac{n}{(1/x_1 + 1/x_i + ... + 1/x_n)}$$

Obviously, Q_{-1} is the harmonic mean for the variate X.

(c) Taking logarithms of both sides of (1) we obtain

$$\ln(Q_p) = \frac{\ln[(x_1^p + x_2^p + ... + x_n^p)/n}{p}$$

When $p \to 0$, $\ln(Q_0)$ is the 0/0 form, hence using L'Hopital Rule

$$\ln(Q_0) = \lim_{p\to 0} \frac{x_1^p \ln x_1 + x_2^p \ln x_2 + ... + x_n^p \ln x_n}{(x_1^p + x_2^p + ... + x_n^p)} = \frac{\ln x_1 + \ln x_2 + ... + \ln x_n}{n}$$

Thus $\ln(Q_0) = \ln(x_1 . x_2, ... x_n)^{1/n}$, $\Rightarrow Q_0 = (x_1 x_2 ... x_n)^{1/n}$, the geometric mean of X.

Conclusion : $Q_1 = A$, $Q_{-1} = H$, $Q_0 = G$.

6*. Let $a = \min\{x_1, ..., x_n\}$, $b = \max\{x_1, ..., x_n\}$; then $r = b - a$.

Now $x_i \geq a$, $\forall_i$, hence $\bar{x} = (\Sigma x_i / n) \geq a$; and so

$$x_i - \bar{x} \leq (x_i - a) \leq (b - a) = r \Rightarrow \Sigma(x_i - \bar{x})^2 \leq nr^2, \quad 1 \leq i \leq n$$

As $(n-1)\hat{S}^2 = nS^2 = \Sigma(x_i - \bar{x})^2$, it follows that

$$(n-1)\hat{S}^2 \leq nr^2 \Rightarrow \hat{S} \leq r[n/(n-1)]^{1/2}; \; nS^2 \leq nr^2 \Rightarrow S \leq r.$$

Chapter 7 : Conditional Expectation

Sec. 7-50. Page 278

1*. Marginal density : $f_2(y) = \int_0^\infty y^{-1} e^{-y} e^{-x/y} dx = e^{-y}, \; 0 < y < \infty.$

$$\therefore \qquad f(x \mid y) = \frac{f(x, y)}{f_2(y)} = \frac{e^{-x/y}}{y}.$$

So $$E(X \mid Y = y) = \int_0^\infty x^{2-1} \cdot \frac{e^{-x/y}}{y} dx = \frac{1}{y} \cdot \frac{\Gamma(2)}{(1/y^2)} = y.$$

2*. Let $f(x, y) = xe^{-x(1+y)}, \; x \geq 0, \; y \geq 0; \quad f(x, y) = 0$, otherwise.

$$f_X(x) = xe^{-x} \int_0^\infty e^{-xy} dy = e^{-x}, \quad x \geq 0. \qquad \left[f(y \mid x) = \frac{f(x, y)}{f_1(x)} = xe^{-xy}\right]$$

$$f_Y(y) = \int_0^\infty x^{2-1} e^{-x(1+y)} dx = \frac{\Gamma(2)}{(1+y)^2} = \frac{1}{(1+y)^2}, \quad y \geq 0.$$

$\therefore \quad E(Y) = \int_0^\infty \frac{y}{(1+y)^2} dy \to \infty$, [divergent integral] i.e. $E(Y)$ does not exist.

$$E(XY) = \int_0^\infty \int_0^\infty xy \cdot xe^{-x(1+y)} dx\, dy = \int_0^\infty x^2 e^{-x} dx \int_0^\infty ye^{-xy} dy = \int_0^\infty x^2 e^{-x} \frac{\Gamma(2)}{x^2} dx = \int_0^\infty e^{-x} dx = 1$$

$$E(Y \mid x) = \int_0^\infty y \cdot f(y \mid x) dy = x \int_0^\infty ye^{-xy} dy = x \cdot \left(\frac{\Gamma(2)}{x^2}\right) = \frac{1}{x}.$$

Thus, both $E(XY)$ and $E(Y \mid x)$ exist, although $E(Y) \to \infty$.

3*. We write $E(Y \mid X) = \mu_{Y \mid X}$, and use conditioning :

$$E\{[Y - g(X)]^2 \mid X\} = E\{[(Y - \mu_{Y \mid X}) + (\mu_{Y \mid X} - g(X)]^2 \mid X\}. \text{ [Use Binomial expansion]}$$

$$= E\{(Y - \mu_{Y \mid X})^2 \mid X\} + 2E\{(Y - \mu_{Y \mid X})(\mu_{Y \mid X} - g(X)) \mid X\} + E\{[\mu_{Y \mid X} - g(X)]^2 \mid X\}.$$

$$\because \quad E\{(Y - \mu_{Y \mid X})(\mu_{Y \mid X} - g(X) \mid X\} = \{[\mu_{Y \mid X} - g(X)] \mid X\} E[Y - E(Y \mid X)] = \varphi(X)[E(Y) - E(Y)] = 0$$

$$\therefore \quad E\{[Y - g(X)]^2 \mid X\} \geq E\{(Y - \mu_{Y \mid X})^2 \mid X\}. \qquad ...(2)$$

Taking expectations of both sides of inequality (2) yields result (1).

4*. Let X be the number of random digits to be generated *until* 3 consecutive zeros are secured. Let Y be the number of random digits to be generated until the first non-zero digit is obtained. Then by Double-E Rule.

$$\begin{aligned}
\mathrm{E}(X) &= \mathrm{E}\{\mathrm{E}(X \mid Y)\} = \mathrm{E}[\phi(Y)] = \sum_{j=1}^{\infty} \phi(Y=j)(P \mid Y=j) \\
&= \sum_{j=1}^{3} \phi(Y=j)\, P(Y=j) + \sum_{j=4}^{\infty} \phi(Y=j)\, P(Y=j) \\
&= \sum_{j=1}^{3} [j + \mathrm{E}(X)]\, q^{j-1} p + \sum_{j=4}^{\infty} 3(q)^{j-1} p \quad \left[Y \sim \text{gem}(p),\ p = \frac{9}{10}\right]
\end{aligned}$$

Now $\quad 3p \sum_{j=4}^{\infty} q^{j-1} = 3pq^3 \dfrac{1}{1-q} = 3q^3 = 0.003, \ \sum_{1}^{3} pq^{j-1} = p(1+q+q^2) = 0.999$

$$p\left\{\sum_{j=1}^{3} j\, q^{j-1}\right\} = p\{1 + 2q + 3q^2\} = 0.9\,(1 + 0.2 + 0.03) = 1.107$$

$\therefore \quad \mathrm{E}(X) = 1.107 + 0.999\, \mathrm{E}(X) + 0.003 \quad \Rightarrow \quad \mathrm{E}(X) = 1110.$

5*. We must determine $f_X(x)$:

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y)\, dy = \int_{x^2}^{x} 6\, dy = 6(x - x^2), \quad 0 \le x \le 1$$

$$\therefore \quad f(y \mid x) = \frac{f(x, y)}{f_X(x)} = \frac{1}{x - x^2}, \ x^2 \le y < x, \quad 0 \le x < 1.$$

$$\mathrm{E}(Y \mid X = a) = \int_{x^2}^{x} f(y \mid a) . y\, dy = \int_{a^2}^{a} y \frac{1}{a - a^2}\, dy = \frac{a^2 - a^4}{2(a - a^2)} = \frac{a(a+1)}{2}.$$

$$\mathrm{E}(Y) = \mathrm{E}[\mathrm{E}(Y \mid X)] = \mathrm{E}\left[\frac{1}{2} X (X+1)\right] = \int_0^1 6x(1-x) . \frac{1}{2} x(x+1)\, dx = 3\int_0^1 x^2(1 - x^2)\, dx = \frac{2}{5}.$$

6*. (i) Conditioning on the values of Y yields

$$\begin{aligned}
p &= P(X < Y) = \int_{-\infty}^{\infty} P\{(X < Y) \mid Y = y\}\ f_Y(y)\, dy = \int_{-\infty}^{\infty} P(X < y)\, f_Y(y)\, dy \\
&= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} f_X(x)\, dx\right) f_Y(y)\, dy = \int_{-\infty}^{\infty} F_X(y)\, f_Y(y)\, dy.
\end{aligned}$$

(ii) $P\{X + Y < a\} = \int_{-\infty}^{\infty} P\{(X+Y) < a \mid Y = y\} f_Y(y)\, dy = \int_{-\infty}^{\infty} P(X < a - y)\, f_Y(y)\, dy = \int_{-\infty}^{\infty} F_X(a-y) f_Y(y)\, dy.$

7*. We enumerate the values of X and Y and more that probability $p = 1/16$ for each outcome $\omega = (x, y)$.

ω	x	y	P
(1, 1)	1	1	p
(1, 2)	1	2	p
(1, 3)	1	3	p
(1, 4)	1	4	p

ω	x	y	P
(2, 1)	2	2	p
(2, 2)	2	2	p
(2, 3)	2	3	p
(2, 4)	2	4	p

ω	x	y	P
(3, 1)	3	3	p
(3, 2)	3	3	p
(3, 3)	3	3	p
(3, 4)	3	4	p

ω	x	y	P
(4, 1)	4	4	p
(4, 2)	4	4	p
(4, 3)	4	4	p
(4, 4)	4	4	p

$y \rightarrow$ / $x \downarrow$	1	2	3	4	$P(x)$
1	p	p	p	p	$4p$
2	0	$2p$	p	p	$4p$
3	0	0	$3p$	p	$4p$
4	0	0	0	$4p$	$4p$
$P(y)$	p	$3p$	$5p$	$7p$	$16p = 1$

(A)

$y \rightarrow$ / $x \downarrow$	1	2	3	4	$P(x)$
1	p	p	$p-\varepsilon$	$p+\varepsilon$	$4p$
2	0	$2p$	$p+\varepsilon$	$p-\varepsilon$	$4p$
3	0	0	$3p$	p	$4p$
4	0	0	0	$4p$	$4p$
$P(y)$	p	$3p$	$5p$	$7p$	$16p$

(B)

(a) Table (A) reads the joint density of (X, Y). Let $P(X = x) = f_1(x)$;

then, $$f_1(1) = f_1(2) = f_1(3) = f_1(4) = 4p = \tfrac{1}{4}.$$

Similarly, $P(Y = y) = f_2(y), f_2(1) = p, f_2(2) = 3p, f_2(3) = 5p, f_2(4) = 7p$ (marginal totals).
Since $f(2, 2) \neq f_1(2) \cdot f_2(2)$, it follows that X, Y are not independent.

(b) Let $0 < \varepsilon < p$. Then, the joint density $g(x, y)$ shown in table (B) has the same f_X and f_Y as in part (a). We thus note that the joint density is not uniquely determined from the knowledge of the marginals.

(c) $P(X \leq 2, Y \leq 3) = f(1, 1) + f(1, 2) + f(1, 3) + f(2, 1) + f(2, 2) + f(2, 3) = 6p = 6/16.$

(d) $\mathsf{E}(X) = 4p(1 + 2 + 3 + 4) = 40p = 5/2$, $\mathsf{E}(X^2) = 4p(1 + 4 + 9 + 16) = 120p = 15/2$,
$\mathsf{E}(Y) = p(1 + 6 + 15 + 28) = 50p = 25/8$, $\mathsf{E}(Y^2) = p(1 + 12 + 45 + 112) = 70p = 85/8$,
$\mathsf{E}(XY) = p[(1 + 2 + 3 + 4) + (8 + 6 + 8) + (27 + 12) + 64] = 135p = 135/16.$

$\therefore$ Var $(X) = 5/4$, Var $(Y) = 55/64$, Cov $(X, Y) = 5/8$, Corr $(X, Y) = (5/8)/\sqrt{5/4 \,.\, 55/64} = 2/\sqrt{11}.$

(e) Observe : $f(y \mid X = 2) = f(x, y)/f_1(2) = 4f(2, y).$

So, $f(1 \mid 2) = 0, \quad f(2 \mid 2) = 8p, \quad f(3 \mid 2) = 4p, \quad f(4 \mid 2) = 4p.$

$$\mathsf{E}(Y \mid X = 2) = \textstyle\sum y f(y \mid 2) = 16p + 12p + 16p = 44p = 11/4.$$

8*. For discrete variates X and Y, $\mathsf{E}(Y \mid X = x) = \sum_y y f_Y(y \mid x) = \sum_y y f(y) = \mathsf{E}(Y)$, since by independence $f_{Y \mid X}(y \mid x) = f_Y(y)$. If variates are continuous,

$$\mathsf{E}(Y \mid X = x) = \int_{-\infty}^{\infty} y f(y \mid x)\, dy = \int_{-\infty}^{\infty} y f_Y(y) = \mathsf{E}(Y), \text{ as before.}$$

Converse. Let X and Y have joint p.d.f. $f(x, y) = 1/\pi, \quad x^2 + y^2 \leq 1$
The marginal densities of X and Y are

$$f_X(x) = 2\sqrt{1 - x^2/\pi}, \; -1 \leq x \leq 1 \; ; \; f_Y(y) = 2\sqrt{1 - y^2/\pi}, \; -1 \leq y \leq 1.$$

Obviously, $f(x, y) \neq f_X(x)\, f_Y(y) \Rightarrow X$ and Y are not independent.

Now $$f(y \mid x) = \frac{f(x, y)}{f_X(x)} = \frac{1}{2\sqrt{1 - x^2}}, \qquad -\sqrt{1 - x^2} < y < \sqrt{1 - x^2}.$$

$\therefore \quad E(Y \mid x) = \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} y f(y \mid x)\, dy = \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \frac{y\, dy}{2\sqrt{1-x^2}} = 0.$ [Odd integrand]

Further, $E(Y) = \int_{-1}^{1} y f_Y(y)\, dy = \frac{2}{\pi}\int_{-1}^{1} y\sqrt{1-y^2}\, dy = 0.$ [Odd integrand]

Thus, $E(Y \mid X = x) = 0 = E(Y)$, although X and Y are not independent. Thus, if X and Y are independent then $E(X \mid Y) = E(X)$, $E(Y \mid X) = E(Y)$. But the converse need not be true.

9*. By normality, $1 = \sum_{i,j} p_{ij} = 10p \Rightarrow p = \frac{1}{10} = 0.1$.

The row totals and column totals in the adjoining table provide the marginal distributions $P(x)$ and $P(y)$ respectively. Thus

X	:	-1	0	1	Y	:	-1	0	1
$P(x)$	:	$2p$	$4p$	$4p$	$P(x)$	:	$2p$	$6p$	$2p$

$\therefore \quad E(X) = -2p + 0 + 4p = 2p = 0.2; \quad E(Y) = -2p + 0 + 2p = 0$

$E(X^2) = 2p + 0 + 4p = 6p; \quad E(Y^2) = 2p + 0 + 2p = 4p.$

$\therefore$ Var $(X) = E(X^2) - E^2(X) = 6p - 4p^2 = 0.6 - 0.04 = 0.56$; Var $(Y) = E(Y^2) - E^2(Y) = 4p - 0 = 0.4$.

$E(XY) = -p + p = 0$ (From table, omitting zero terms, seven terms contribute zeros)

Since Cov $(X, Y) = E(XY) - E(X)\, E(Y) = 0$, X and Y are uncorrelated. However

$P(X = 0, Y = 0) = 2p$, $P(X = 0) = 4p$, $P(Y = 0) = 6p$; so $P(0, 0) \neq P_X(0)\, .\, P_Y(0)$.

Thus X and Y are not independent.

$P(X = x \mid Y = 0) = P(x, 0)/P(Y = 0) = P(x, 0)/6p$

$\therefore \quad P(X = -1 \mid Y = 0) = P(-1, 0)/6p = 2p/6p = \frac{1}{3}$

$P(X = 0 \mid Y = 0) = P(0, 0)/6p = 2p/6p = \frac{1}{3}$; $P(X = 1 \mid Y = 0) = P(1, 0)/6p = 2p/6p = \frac{1}{3}$

Again $\quad f(Y \mid x) = f(x, y)/f_X(x)$; so when $x = -1$, $f(y \mid x = -1) = f(-1, y)/2p$.

$\therefore \quad E(Y \mid X = -1) = \sum y\, f(y \mid x = -1) = [\sum y\, f(-1, y)]/2p = [0 + 0 + 0]/2p = 0$

$E(Y^2 \mid X = -1) = [\sum y^2 f(-1, y)]/2p = [0 + 0 + 0]/2p = 0$

Thus, $\quad$ Var $(Y \mid X = -1) = E(Y^2 \mid X = -1) - [E(Y \mid X = -1)]^2 = 0$.

10*. Firstly we check up normality :

$$\int_0^{\infty} f(x, y)\, dy = e^{-x}\int_0^{x}(1 - e^{-y})\, dy + (1 - e^{-x})\int_x^{\infty} e^{-y}\, dy = e^{-x}(x + e^{-x} - 1) + e^{-x}(1 - e^{-x}) = xe^{-x} \quad ...(1)$$

$$\therefore \quad \int_0^{\infty}\int_0^{\infty} f(x, y)\, dy\, dx = \int_0^{\infty} x^{2-1} e^{-x} dx = \Gamma(2) = 1.$$

This shows that $f(x, y)$ is a bonafide density. As a by product, (1) yields the marginal density :

$$f_1(x) = xe^{-x}, \quad 0 < x < \infty \quad ...(1)$$

By symmetry : $\quad f_2(y) = ye^{-y}, \; 0 < y < \infty \quad ...(2)$

The joint density $g(x, y) = f_1(x)\, .\, f_2(y)$, $\quad 0 \leq x, y < \infty$, has the same marginals $f_1(x)$ and $f_2(y)$ as does $f(x, y)$. However, $f(x, y) \neq g(x, y)$.

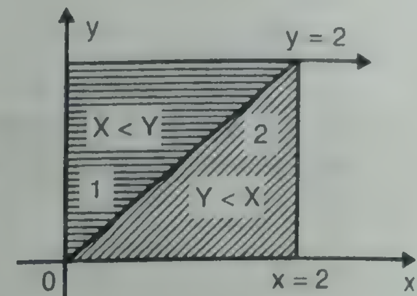
(a) $f(y \mid x) = f(x, y)/f_1(x) = e^x \cdot f(x, y)/x$.

$$\therefore \quad \mathrm{E}(Y \mid x) = \frac{(e^x - 1)}{x}\int_x^{\infty} ye^{-y}\,dy + \frac{1}{x}\int_0^x y(1-e^{-y})\,dy$$

$$= \left(\frac{e^x - 1}{x}\right)\left[e^{-y}(1+y)\right]_{\infty}^{x} + \frac{1}{x}\left[\frac{1}{2}y^2 + e^{-y}(1+y)\right]_0^x = 1 + \left(\frac{x}{2}\right).$$

(b) $p = P\{X \le 2, Y \le 2\}$

$$= \int_{x=0}^{2}\int_{y=x}^{2} f(x,y)\,dx\,dy + \int_0^2\int_{y=0}^{x} f(x,y)\,dx\,dy$$



$$\therefore \quad p = \int_0^2 (1-e^{-x})\,dx \int_x^2 e^{-y}\,dy + \int_0^2 e^{-x}\,dx \int_0^x (1-e^{-y})\,dy$$

$$= \int_0^2 (1-e^{-x})(e^{-x}-e^{-2})\,dx + \int_0^2 e^{-x}(x+e^{-x}-1)\,dx$$

$$= \int_0^2 [(x+e^{-2})e^{-x} - e^{-2}]\,dx = (1 - 4e^{-2} - e^{-4}).$$

(c) $\mathrm{E}(X) = \int_0^{\infty} (x.e^{-x})\,x\,dx = \Gamma(3) = 2;\ \mathrm{E}(X^2) = \int_0^{\infty} (x.e^{-x})\,x^2\,dx = \Gamma(4) = 6$

$\therefore$ Var $(X) = \mathrm{E}(X^2) - \mathrm{E}^2(X) = 6 - 4 = 2.$ Var $(Y) = 2$, by symmetry.

$$\mathrm{E}(XY) = \int_0^{\infty} ye^{-y}\,dy \int_0^y x(1-e^{-x})\,dx + \int_0^{\infty} xe^{-x}\,dx \int_0^x y(1-e^{-y})\,dy$$

$$= 2\int_0^{\infty} ye^{-y}\,dy \int_0^y x(1-e^{-x})\,dx \qquad \text{[2nd integral = 1st integral]}$$

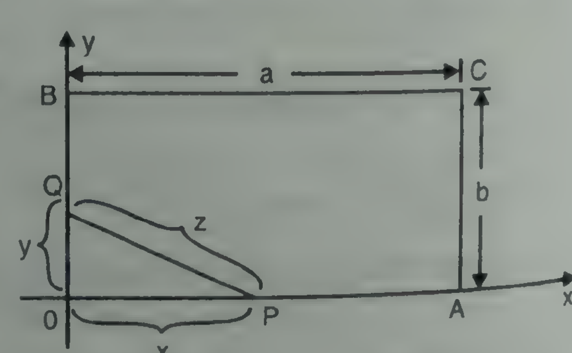
$$= 2\int_0^{\infty} ye^{-y}\left[\frac{1}{2}y^2 + e^{-y}(1+y) - 1\right]dy = \int_0^{\infty} e^{-y}[y^3 + 2(y+y^2)e^{-y} - 2y]\,dy$$

$$= \Gamma(4) + 2.\left(\frac{1}{4}\right).\Gamma(2) + 2.\left(\frac{1}{8}\right).\Gamma(3) - 2.\Gamma(2) = 5.$$

$$\therefore \quad \text{Cov}(X, Y) = \mathrm{E}(XY) - \mathrm{E}(X)\mathrm{E}(Y) = 5 - 2\times 2 = 1 \Rightarrow \rho_{XY} = \frac{\sigma_{XY}}{\sigma_X . \sigma_Y} = \frac{1}{2}$$

Note. $\mathrm{E}[\mathrm{E}(Y \mid x)] = \mathrm{E}\left[1 + \frac{1}{2}X\right] = 1 + \frac{1}{2}\mathrm{E}(X) = 1 + \frac{1}{2}(2) = 2 = \mathrm{E}(X) \Rightarrow \mathrm{E}(X) = \mathrm{E}[\mathrm{E}(Y \mid X)].$

11*. Selection of two random points on the boundary of a Rectangle $a \times b$ gives rise to the following unique disjoint events :



A = {Both points lie on the same side a};

B = {Both points lie on the same side b};

A_1 = {Points lie on opposite sides a};

B_1 = {Points lie on opposite sides b};

C = {Points lie on adjacent sides}.

Let $2s = 2(a + b)$ be the perimeter of the rectangle, then the measure of length available to choose any point on the boundary is $2s$; and measure of length for the specific sides is a or b as the case may be.

$P(A) = 2\left(\frac{a}{2s}\cdot\frac{a}{2s}\right) = \frac{a^2}{2s^2} = P(A_1)\ P(B) = 2\left(\frac{b}{2s}\cdot\frac{b}{2s}\right) = \frac{b^2}{2s^2} = P(B_1)\ P(C) = 8\left(\frac{a}{2s}\cdot\frac{b}{2s}\right) = \frac{2ab}{s^2}.$

If Z is the distance between the two points, then

$$E(Z^2 \mid A) = \int_0^a \int_0^a f(x_1, x_2)(x_1 - x_2)^2\, dx_1\, dx_2 = \frac{1}{a^2}\int_0^a \int_0^a (x_1 - x_2)^2\, dx_1\, dx_2 = \frac{a^2}{6} \qquad \ldots(1)$$

$$E(Z^2 \mid B) = \frac{b^2}{6}.$$

$$E(Z^2 \mid C) = \frac{1}{ab}\int_0^a \int_0^b (x^2 + y^2)\, dx\, dy = \frac{1}{3}(a^2 + b^2).$$

$$E(Z^2 \mid A_1) = E\{b^2 + (X_1 - X_2)^2\} = b^2 + E(X_1 - X_2)^2 = b^2 + \left(\frac{a^2}{6}\right), \qquad \text{[by (1)]}$$

$$E(Z^2 \mid B_1) = a^2 + \left(\frac{b^2}{6}\right), \ (a \rightleftharpoons b)$$

The total expectation of the variate Z^2, by Total-E Rule is

$$\begin{aligned} E(Z^2) &= \Sigma E(Z^2 \mid A)\, P(A) = \left(\frac{1}{6}s^2\right)[a^4 + 4a^3b + 6a^2b^2 + 4ab^3 + b^4] \\ &= \frac{(a+b)^4}{6s^2} = \frac{s^2}{6} = \frac{(a+b)^2}{6}. \end{aligned}$$

Chapter 8 : *Generating Functions or Transform Methods*

Sec. 8-19. Page 291

1*. If $t = 0$ then $M(t) = E(e^{0X}) = 1$; but $M(t) = t/(1-t) = 0$, at $t = 0$.

Hence, $t/(1 - t)$ cannot be a m.g.f. of any variate X.

2*. $M(t : X) = E(e^{tX}) = \int_{-a}^{a} \frac{e^{tX}}{2a}\, dx = \frac{1}{2at}[e^{at} - e^{-at}] = \frac{\sinh at}{at}.$

The power series for $M(t)$ and sinh at provide

$$\sum_{r=0}^{\infty} \frac{t^r}{r!}\mu'_r = \frac{1}{at}\sum_{r=1}^{\infty} \frac{(at)^{2r+1}}{(2r+1)!} = \sum_{r=0}^{\infty} \frac{(at)^{2r}}{(2r)!}\,\frac{1}{(2r+1)}. \qquad \ldots(1)$$

Obviously, $\mu = E(X) =$ Coeff. of $t = 0$. This shows that the simple and central moments are identical. Further, there are no odd powered terms it t, hence $\mu_{2n+1} = 0$. Equating coefficient of $t^{2n}/(2n)$! on both sides in (1) provides $\mu_{2n}/(2n + 1)$.

3*. Here $F(x) = 0$, when $x < 0$, $F(0) = \frac{1}{2}$; hence $P(X = 0) = \frac{1}{2}$.

For $x > 0$, $f(x) = F'(x) = \frac{1}{2}e^{-x}$. The Dirichlet's form for m.g.f. is thus

$$M_X(t) = E(e^{tX}) = \tfrac{1}{2}e^{t} + \int_0^{\infty} \tfrac{1}{2}e^{-x}e^{tx}\, dx = \frac{1}{2} + \frac{1}{2}\,\frac{1}{1-t} = \frac{2-t}{2(1-t)}, \quad t < 1.$$

4*. $M(t)=\sum_{r=0}^{\infty}\frac{t^r}{r!}\mu_r'=1+\sum_{r=1}^{\infty}\frac{t^r}{r!}(0.6)=(0.4)+(0.6)\sum_{r=0}^{\infty}\frac{t^r}{r!}=0.4+(0.6)e^t$

Since $M(t)=\Sigma P(X=k)\,e^{tk}=f(0)+f(1)\,e^t+f(2)\,e^{2t}+\ldots$, 1, 2, ..., it follows that

$$P(X=0)=0.4,\ P(X=1)=0.6,\ P(X\geq 2)=0.$$

5*. Twice differentiating $M(t)$ w.r.t. parameter t, we get

$M'(t)=cm'(t)\,e^{c(m-1)},\ M''(t)=ce^{c(m-1)}[m''(t)+cm'^2(t)].$

Now $\mathsf{E}(Y)=M'(0), \mathsf{E}(Y^2)=M''(0),\ \mathsf{E}(X)=m'(0), \mathsf{E}(X^2)=m''(0),\ m(0)=1.$

$\therefore \mathsf{E}(Y)=c\mu, \mathsf{E}(Y^2)=c[\mathsf{E}(X^2)+c\mu^2]=c[\sigma^2+\mu^2]+c^2\mu^2 \Rightarrow \mathrm{Var}(Y)=\mathsf{E}(Y^2)-c^2\mu^2=c(\sigma^2+\mu^2).$

Aliter. Using exponential series we get

$$M(t)=e^{c[\mu t+\frac{1}{2}\mu_2' t^2+\ldots]}=1+c[\mu t+\tfrac{1}{2}\mu_2' t^2+\ldots]+\tfrac{1}{2}c^2[\mu t+\tfrac{1}{2}\mu_2' t^2+\ldots]^2+\ldots$$

$\mathsf{E}(Y)$ = Coeff. of $t=c\mu$; $\mathsf{E}(Y^2)$ = coeff. of $\frac{1}{2}t^2=c\mu_2'+c^2\mu^2$

$\therefore\quad$ Var (Y) = $\mathsf{E}(Y^2)-\mathsf{E}^2(Y)=c\mu_2'=c(\sigma^2+\mu^2).$

6*. $M(t:X)=\mathsf{E}(e^{tx})=\int_0^{\infty}f(x)\,e^{tx}dx+\sum_0^{\infty}p(x)\,e^{tx}$

$$M(t)=\frac{1}{2\lambda}\int_0^{\infty}e^{-(1-\lambda t)x/\lambda}x^{1-1}dx+\frac{e^{-\lambda}}{2}\sum_{x=0}^{\infty}\frac{(\lambda e^t)^x}{x!}. \qquad \left[\sum_0^{\infty}\frac{\theta^k}{k!}=e^{\theta}\right]$$

$$=\tfrac{1}{2}(1-\lambda t)^{-1}+\tfrac{1}{2}e^{-\lambda}.e^{\lambda e t}=\tfrac{1}{2}(1-\lambda t)^{-1}+\tfrac{1}{2}e^{\lambda(t+t^2/2+\ldots)} \qquad \ldots(1)$$

To find Var (X), we utilize series expansions to obtain

$$M(t)=\tfrac{1}{2}(1+\lambda t+\lambda^2t^2+\ldots)+\tfrac{1}{2}[1+\lambda(t+\tfrac{1}{2}t^2+\ldots)+\tfrac{1}{2}\lambda^2(t+\tfrac{1}{2}t^2+\ldots)^2+\ldots]$$

$\therefore\quad \mu$ = Coeff. of $t=\lambda$; $\mu_2'=\mathsf{E}(X^2)$ = Coeff. of $\frac{1}{2}t^2=(3\lambda^2/2)+(\lambda/2)$

$\therefore\quad$ Var (X) = $\mu_2'-(\mu_1')^2=\frac{1}{2}(\lambda^2+\lambda).$

Note. $M(t)=\frac{1}{2}$ m.g.f. Expo $(\lambda^{-1})+\frac{1}{2}$ m.g.f. Pois (λ).

7*. By Multi-Stage p-Rule, we have $f(x)=p\{e^{-\frac{1}{2}x^2}/\sqrt{2\pi}\}+q\{e^{-\frac{1}{2}(x-1)^2}/\sqrt{2\pi}\}.$

$$\therefore\quad M(t:X)=\mathsf{E}(e^{tx})=p\int_{-\infty}^{\infty}e^{tx}\frac{e^{-\frac{1}{2}x^2}}{\sqrt{2\pi}}dx+q\int_{-\infty}^{\infty}e^{tx}\frac{e^{-\frac{1}{2}(x-1)^2}}{\sqrt{2\pi}}dx$$

$$=p[\text{m.g.f. N}(0,\,1)]+q[\text{m.g.f. N}(1,\,1)]=pe^{t^2/2}+qe^{(t+t^2/2)}.$$

To find Var (X), we use exponential expansions. Thus

$$M(t)=p[1+\tfrac{1}{2}t^2+\ldots]+q[1+(t+\tfrac{1}{2}t^2)+\tfrac{1}{2}(t^2+t^3+\tfrac{1}{4}t^4)+\ldots]$$

$\therefore\quad \mathsf{E}(X)$ = Coeff. of $t=q, \mathsf{E}(X^2)$ = Coeff. of $\frac{1}{2}t^2=p+2q.$

$$\mathrm{Var}(X)=\mathsf{E}(X^2)-\mathsf{E}^2(X)=p+2q-q^2=1+pq.$$

Sec. 8-46. Page 302

1*. We find y_0 by Normality : $\Sigma f(x) = 1$; so

$$1 = y_0 \int_0^\infty e^{-x/\sigma} = \frac{y_0}{(1/\sigma)} = \sigma y_0 \Rightarrow y_0 = 1/\sigma.$$

To find quartiles, we use definition to secure

$$\frac{r}{4} = \frac{1}{\sigma} \int_0^{Q_r} e^{-x/\sigma} dx = 1 - e^{-Q_r/\sigma} \Rightarrow Q_r = -\sigma \log\left(1 - \frac{r}{4}\right).$$

So, $$Q_3 = \sigma \log 4, Q_1 = \sigma \log(4/3), Q = Q_3 - Q_1 = \sigma \log 3.$$

To find k_r we first determine m.g.f. :

$$M(t) = \mathsf{E}(e^{tX}) = \int_0^\infty \frac{e^{tx} e^{-x/\sigma}}{\sigma} dx = \frac{1}{\sigma} \int_0^\infty e^{-(\sigma^{-1} - t)X} dx = \frac{1}{\sigma} \cdot \frac{\sigma}{1 - \sigma t} = (1 - \sigma t)^{-1}, (|t| < 1/\sigma) \quad \ldots(1)$$

$$= \sum_{r=0}^\infty (\sigma t)^r = \sum_{r=0}^\infty \frac{t^r}{r!}(r!\sigma^r) = \sum_{r=0}^\infty \mu'_r \frac{t^r}{r!} \qquad \text{[by Bin. expansion]}$$

Thus, $\mu'_r = r!\sigma^r$.

Also, $K(t) = \ln M(t) = -\ln(1 - \sigma t) = \sum_{r=1}^\infty \frac{(\sigma t)^r}{r} = \sum_{r=1}^\infty \frac{t^r}{r!}[(r-1)!\sigma^r] = \sum_{r=1}^\infty \frac{t^r}{r!} k_r.$

$$\therefore \qquad k_r = (r-1)!\,\sigma_r$$

This gives, $k_1 = \sigma, k_2 = \mu_2 = \sigma^2, k_3 = \mu_3 = 2\sigma^3, k_4 = 6\sigma^4$

So $\beta_1 = \dfrac{\mu_3^2}{\mu_2^3} = \dfrac{4\sigma^6}{\sigma^6} = 4;$ $\qquad \beta_2 = \dfrac{\mu_4}{\mu_2^2} = \dfrac{k_4 + 3k_2^2}{k_2^2} = 9.$ $\qquad [\beta_2 = \alpha_4, \beta_1 = \alpha_3^2]$

2*. Since $\mathsf{E}(X) = \mu = 0$, given, $\mu'_n = \mu_n \cdot \mu_2 = k_2, \mu_3 = k_3$, etc. Now

$$M(t : X^2) = \mathsf{E}(e^{tX^2}) = \mathsf{E}\left\{\sum_{r=0}^\infty \frac{t^r X^{2r}}{r!}\right\} = \sum_{r=0}^\infty \frac{t^r}{r!} \mathsf{E}^{(2r)} = \sum_{r=0}^\infty \frac{t^r \mu_{2r}}{r!} \qquad \text{[Lin E]}$$

So $\quad K(t : Y) = \ln M(t : X^2) = \ln\{1 + \Sigma(t^r \mu_{2r}/r!)\}, 1 \le r < \infty.$

$$\therefore \quad \sum_{r=1}^\infty \lambda_r \frac{t^r}{r!} = \left[\sum_{r=1}^\infty \frac{t^r}{r!}\mu_{2r}\right] - \frac{1}{2}\left[\sum_{r=1}^\infty \frac{t^r}{r!}\mu_{2r}\right]^2 + \frac{1}{3}\left[\sum_{r=1}^\infty \frac{t^r}{r!}\mu_{2r}\right]^2 - \ldots$$

Equating the coefficients of t, t^2, t^3, we get

$$\lambda_1 = \mu_2, \qquad \lambda_2 = \mu_4 - \mu_2^2, \qquad \lambda_3 = \mu_6 - 3\mu_2\mu_4 + 2(\mu_2)^3 \qquad \ldots(1)$$

Recall : $\mu_2 = k_2, \quad \mu_4 = k_4 = 3k_2^2, \quad \mu_5 = k_5 + 10k_3k_2, \quad \mu_6 = k_6 + 15k_4k_2 + 10k_3^2 + 15k_2^3 \quad$ [§7-42]

$$\therefore \quad \lambda_1 = k_2, \; \lambda_2 = (k_4 + 3k_2^2) - k_2^2 = k_4 + 2k_2^2$$

$\lambda_3 = (k_6 + 15k_4k_2 + 10k_3^2 + 15k_2^3) - 2k_2(k_4 + 3k_2^2) + 2k_2^3 = k_6 + 2(5k_3^2 + 6k_2k_4) + 8k_2^3.$ [Putting in (1)]

3*. We twice differentiate w.r.t. t the given relation to obtain :

$$M'(t:X+Y)=M_X(t)\,M_Y'(t)+M_X'(t)\,M_Y(t)$$

$$M''(t:X+Y)=M_X(t)\,M_Y''(t)+2M_X'(t)\,M_Y'(t)+M_X''(t)\,M_Y(t)$$

Setting $t=0$, using $M_Z(0)=1$, $M_Z'(0)=\mathrm{E}(Z)$, $M''_Z(0)=\mathrm{E}(Z^2)$, we get

$$\mathrm{E}(X+Y)^2=\mathrm{E}(X^2)+2\mathrm{E}(X)\mathrm{E}(Y)+\mathrm{E}(Y^2)$$

$$\therefore\quad \mathrm{E}(X^2)+\mathrm{E}(Y^2)+2\mathrm{E}(XY)=\mathrm{E}(X^2)+2\mathrm{E}(X)\mathrm{E}(Y)+\mathrm{E}(Y^2)\ \Rightarrow\ \mathrm{E}(XY)=\mathrm{E}(X)\,.\,\mathrm{E}(Y).$$

This gives Cov $(X, Y)=0\ \Rightarrow X$ and Y are uncorrelated.

Comments. $M^{(k)}(t:Z)=D^k\,\mathrm{E}(e^{tZ})=\mathrm{E}(D^k e^{tZ})=\mathrm{E}(Z^k\,.\,e^{tZ})$

The DUES by t is justified by the regularity of exponential function and the simple structure of the c.d.f. F_X. Upon setting $t=0$, we get $\mu_k'=M_Z^{(k)}(0)$.

4*. We shall use m.f.g. technique. Now

$$M(t:X_j-\mu)=e^{-\mu t}M(t:X_j)\Rightarrow M(t:X_j)=e^{\mu t}M(t:X_j-\mu)$$

$$M(t:\overline{X})=[M(t/n:X_j)]^n=e^{\mu t}[M(t/n:X_j-\mu)]=e^{\mu t}\left[1+\sum_{r=2}^{\infty}\frac{(t/n)^r}{r!}\mu_r\right]^n\ \text{[m.g.f. expansion]}$$

$$k(t:\overline{X})=\ln M(t:\overline{X})=\mu t+n\ln\left[1+\Sigma(t^r/r!n^r)\mu_r\right]$$

$$=\mu t+n\left[\sum_{r=2}^{\infty}\frac{t^r}{r!}\frac{\mu_r}{n^r}-\frac{1}{2}\left(\sum_{2}^{\infty}\frac{t^r}{r!}\frac{\mu_r}{n^r}\right)^2+\frac{1}{3}\left(\sum_{2}^{\infty}\frac{t^r}{r!}\frac{\mu_r}{r^r}\right)^3\cdots\right]$$

Equating Coeff. of t^r/r ! on both sides ($r=1, 2, 3, 4$)

$$k_1(\bar{x})=\mu,\ k_2(\bar{x})=\mu_2(\bar{x})=\mu_2/n,\ k_3(\bar{x})=\mu_3(\bar{x})=\mu_3/n^2$$

$$k_4(\bar{x})=\frac{\mu_4}{n^3}-\frac{3\mu_2^2}{n^3}\Rightarrow\mu_4(\bar{x})=k_4(\bar{x})+3k_2^2(\bar{x})=\frac{\mu_4-3\mu_2^2}{n^3}+\frac{3\mu_2^2}{n^2}\qquad\ldots(1)$$

These are the first four central moments of $\overline{X}$ (about μ) in terms of the central moments of variables X_j. Now :

$$(n-1)S^2\equiv\Sigma(X_j-\overline{X})^2=\Sigma(X_j-\mu)^2-n(\overline{X}-\mu)^2=\Sigma(X_j-\mu)^2-(1/n)[\Sigma(X_j-\mu)]^2\qquad\ldots(2)$$

$$=(\Sigma Y_j)^2-(\Sigma Y_j)^2/n.\qquad[Y_i=X_j-\mu]\qquad\ldots(3)$$

Take expected value in (2) and use (2) to get

$$(n-1)\,\mathrm{E}(S^2)=n\mathrm{E}(X_j-\mu)^2-n\mathrm{E}(\overline{X}-\mu)^2=n\mu_2-n(\mu_2/n)=(n-1)\mu_2$$

Thus, $$\mathrm{E}(S^2)=\mu_2=\sigma^2\qquad\ldots(4)$$

Square (2) or (3) to obtain :

$$(n-1)^2S^4=[\Sigma(X_j-\mu)^2]^2-2n(\overline{X}-\mu)^2\,\Sigma(X_j-\mu)^2+n^2(\overline{X}-\mu)^4$$

$$= (\Sigma Y_j^2)^2 - (2/n)(\Sigma Y_j)^2(\Sigma Y_j^2) + n^2(\bar{X}-\mu)^4$$

$$= [\Sigma Y_j^4 + \sum_{i \neq j} Y_i^2 Y_j^2] - (2/n)[\Sigma Y_i^2 + \sum_{i \neq j} Y_i Y_j](\Sigma Y_k^2) + n^2(\bar{X}-\mu)^4$$

$$= [1-(2/n)](\Sigma Y_i^4 + \sum_{i \neq j} Y_i^2 Y_j^2) - (2/n)(\sum_{i \neq j} Y_i Y_j^3 + \sum_{i \neq j \neq k} Y_i Y_j Y_k^2) + n^2(\bar{X}-\mu)^4.$$

We take expected value of both sides, observe that Y_i and Y_j, etc. are indep. and that $\mathsf{E}(Y_j)=0$. This gives

$$(n-1)^2\,\mathsf{E}(S^4) = [(1-(2/n)][n\mu_4 + n(n-1)(\mathsf{E}Y_j^2)^2] + n^2\mu_4(\bar{x})$$

$$= (n-2)[\mu_4 + (n-1)\mu_2^2] + [(\mu_4 - 3\mu_2^2)/n^{-1} + 3\mu_2^2], \quad \text{[by (1)]}$$

$$= [(n-1)^2/n]\,\mu_4 + [(n-1)/n][n^2-2n+3]\,\mu_2^2$$

$\therefore \quad \mathsf{E}(S^4) = (\mu_4/n) + [(n^2-2n+3)/n(n-1)]\mu_2^2$

$$\text{Var}\,(S^2) = \mathsf{E}(S^4) - \mathsf{E}^2(S^2) = \frac{\mu_4 - 3\mu_2^2}{n} + \frac{2\mu_2^2}{n-1}.$$

(b) $\text{Cov}\,(\bar{X}, S^2) = \mathsf{E}[(\bar{X}-\mu)(S^2-\mu_2)] = \mathsf{E}[(\bar{X}-\mu)S^2 - \mu_2(\bar{X}-\mu)] = \mathsf{E}[(\bar{X}-\mu)S^2].$

Now $(n-1)(\bar{X}-\mu)S^2 = (\bar{X}-\mu)[\Sigma(X_j-\mu)^2 - n(\bar{X}-\mu)^2],$ [by (2)]

$$= (\Sigma Y_j)(\Sigma Y_j^2)/n - n(\bar{X}-\mu)^3, \quad \text{[by (3)]}$$

$$= \left(\sum_{j=1}^{n} Y_j^3 + \sum_{i \neq j} Y_i Y_j^2\right)\Big/ n - n(\bar{X}-\mu)^3.$$

Take Expectation, use $\mathsf{E}(Y_i)=0$, use $\mu_3(\bar{x})$ from (1), to get

$$\text{Cov}\,(\bar{X}, S^2) = \frac{n\mu_3}{n(n-1)} + 0 - \frac{n}{n-1}\frac{\mu_3}{n^2} = \frac{\mu_3}{n}.$$

Note. We may also use the following in the above solution.

$$\sum_{ij}(X_i - X_j)^2 = \sum_{i,j}[(X_i-\bar{X}) - (X_j-\bar{X})]^2 = \sum_{i,j}[(X_i-\bar{X})^2 + (X_j-\bar{X})^2 - 2(X_i-\bar{X})(X_j-\bar{X})]$$

$$= n\Sigma(X_i-\bar{X})^2 + n\Sigma(X_j-\bar{X})^2 - 2\Sigma(X_i-\bar{X})\,\Sigma(X_j-\bar{X})] = 2n\Sigma(X_i-\bar{X})^2 = 2n(n-1)\,S^2.$$

Thus, $S^2 = \Sigma\Sigma(X_i - X_j)^2/2n(n-1).$

Sec. 8-61. Page 311

1*. $\mathsf{E}(X^r) = \int_0^\infty \lambda e^{-\lambda x} x^r\,dx = \dfrac{\Gamma(r+1)}{\lambda^r} = \dfrac{r!}{\lambda^r};\ \mathsf{E}(X) = \dfrac{1}{\lambda}, \mathsf{E}(X^2) = \dfrac{2}{\lambda^2}, \text{Var}\,(X) = \dfrac{1}{\lambda^2}.$

$$p_n = P\{nh \le X \le (n+1)h\} = \int_{nh}^{(n+1)h} \lambda e^{-\lambda x}\,dx = e^{-\lambda nh}(1-e^{-\lambda h}) = (1-e^{-2\theta})\,e^{-\lambda nh}, \quad [\lambda h = 2\theta].$$

$$G_Y(t) = \mathrm{E}(t^Y) = (1-e^{-2\theta})\sum_{n=0}^{\infty} e^{-\lambda nh} t^{(2n+1)h/2} = (1-e^{-2\theta})t^{h/2}\sum_{n=0}^{\infty}(e^{-2\theta}t^h)^n$$

$$= (1-e^{-2\theta})\,t^{h/2}.(1-e^{-2\theta}\;t^h)^{-1}.$$

To find moments of Y, we use logarithmic differentiation of $G(t)$.

$$G'(t) = G(t)\left\{\frac{h}{2t} + \frac{ht^{h-1}\,e^{-2\theta}}{1-e^{-2\theta}t^h}\right\} \Rightarrow \mathrm{E}(Y) = G'(1) = \frac{h}{2}\left(\frac{1+e^{-2\theta}}{1-e^{-2\theta}}\right) = \left(\frac{\theta}{\tanh\theta}\right)\frac{1}{\lambda} = \left(\frac{\theta}{\tanh\theta}\right)\mathrm{E}(X).$$

Since $0 < \tanh\theta < \theta$, we readily obtain $\mathrm{E}(Y) > \mathrm{E}(X)$. Further

$$G''(t) = G'(t)\left\{\frac{h}{2t} + \frac{he^{-2\theta}t^{h-1}}{1-e^{-2\theta}t^h}\right\} + G(t)\left\{\frac{-h}{2t^2} + h.e^{-2\theta}\,.\,\frac{(h-1)(1-t^he^{-2\theta})t^{h-2} + he^{-2\theta}t^{2h-2}}{(1-e^{-2\theta}t^h)^2}\right\}$$

$$G''(1) = \frac{h}{2}\coth\theta\left\{\frac{h}{2} + \frac{he^{-2\theta}}{1-e^{-2\theta}}\right\} + \left\{\frac{-h}{2} + \frac{h(h-1+e^{-2\theta})e^{-2\theta}}{(1-e^{-2\theta})^2}\right\} = \frac{h^2}{4}\coth^2\theta + \frac{(e^{-4\theta}+2he^{-2\theta}-1)}{(1-e^{-2\theta})^2}$$

$$\mathrm{Var}(Y) = G''(1) + G'(1) - [G'(1)]^2 = \frac{h^2e^{-2\theta}}{(1-e^{-2\theta})^2} = \frac{(4\theta^2/\lambda^2)}{4[(e^{\theta}-e^{-\theta})/2]^2} = \left(\frac{\theta}{\sinh\theta}\right)^2\frac{1}{\lambda^2}$$

$$= (\theta/\sinh\theta)^2\,\mathrm{Var}(X).$$

Since $\sinh\theta > \theta$, $\mathrm{Var}(Y) = (\theta/\sinh\theta)^2\,\mathrm{Var}(X) \Rightarrow \mathrm{Var}(Y) < \mathrm{Var}(X)$.

2*. Let the ticket bear the number $Y_1Y_2Y_3Y_4Y_5Y_6$. Obviously, Y_j's are i.i.d. variates with $P(Y_j = k) = p = 1/10$, $0 \le k \le 9$. The p.g.f. of Y_j's are indentical, each equal to

$$G(t : Y_j) = (1/10)(1+t+\ldots+t^9) = (1/10)(1-t^{10})/(1-t)$$

So, $G(t : Y_1+Y_2+Y_3) = (1/10)^3(1-t^{10})^3(1-t)^{-3} = G(t : Y_4+Y_5+Y_6) = G(t)$, say

$$P(Y_1+Y_2+Y_3 = r) = \text{Coeff. of } t^r \text{ in } G(t);\;\; P(Y_4+Y_5+Y_6 = r) = \text{Coeff. of } t^{-r} \text{ in } G(t^{-1})$$

$\therefore$ $$p = P(Y_1+Y_2+Y_3 = Y_4+Y_5+Y_6) = \text{Coeff. of } t^0 \text{ in } G(t).G(t^{-1})$$

So $$p = \text{Coeff. of } t^0 \text{ in } (1/10)^6(1-t^{10})/t^{27}.(1-t)^6.$$

Recall ***Neg-bin Thm.*** : $(1-x)^{-6} = \sum\binom{5+r}{5}x^r, \quad r = 0, 1, 2, \ldots$

Now, $(1-t^{10})^6 = 1-\binom{6}{1}t^{10}+\binom{6}{2}t^{20}+\ldots;\;(1-t)^{-6} = 1+\binom{6}{5}t+\binom{7}{5}t^2+\ldots+\binom{5+r}{5}t^r+\ldots$

$$p = \left(\frac{1}{10}\right)^6\left[\binom{32}{5}-\binom{6}{1}\binom{22}{5}+\binom{6}{2}\binom{12}{5}\right] = 0.05525 \qquad [take\ r = 27,\ 17,\ 7]$$

3*. The generating function for one throw is

$$G(t : X_1) = \frac{1}{6}(t+t^2+\ldots+t^6) = \frac{t}{6}\left(\frac{1-t^6}{1-t}\right) = \frac{t}{6}(1-t^6)(1-t)^{-1}.$$

Since the throws are independent, with identical distributions

$$G(t:S) = G(t:X_1+X_2+\dots+X_n) = [G(t:X_1)]^n = t^n(1-t^6)^n(1-t)^{-n}/6^n.$$

$$= \frac{t^n}{6^n}\left[1-\binom{n}{1}t^6+\binom{n}{2}t^{12}-\dots\right]\left[1+\binom{n}{1}t+\binom{n+1}{2}t^2+\binom{n+1}{3}t^3+\dots+\binom{n+k-1}{k}t^k\dots\right] \quad \dots(1)$$

$$\therefore \qquad P(S=n+5) = \text{Coeff. of } t^{n+5} \text{ in (1)} = \binom{n+4}{5}\left(\frac{1}{6}\right)^n$$

$$P(S=n+4) = \text{Coeff. of } t^{n+4} \text{ in (1)} = \binom{n+3}{4}\left(\frac{1}{6}\right)^n$$

Comments. We juxtapose the infinite series in (1)

$$G(t:S) = \frac{t^n}{6^n}\sum_{j=0}^{n}\binom{n}{j}(-t^6)^j\sum_{k=0}^{\infty}\binom{n+k-1}{k}t^k = \sum_{j=0}^{n}\sum_{k=0}^{\infty}\binom{n}{j}\binom{n+k-1}{k}\frac{(-1)^j}{(6)^n}(t)^{n+k+6j} \quad \dots(A)$$

To find $P(X=r)$, we put $6j+n+k=r$, and (A) gives

$$P(S=r) = \frac{1}{6^n}\sum_{j=0}^{n}(-1)^j\binom{n}{j}\binom{r-6j-1}{n-1}, \qquad \left[\because \binom{n}{r}=\binom{n}{n-r}\right]$$

4*. Let $G(u) = \Sigma p_k u^k, (0 \le k < \infty)$ be the p.g.f. for the distribution. We multiply (1) by u^k and sum for all values of k, $[\Sigma u^k p_{k-1} = uG(u)]$ taking $p_k = 0$, for $k < 0$. This gives

$$dG(u)/dt = \lambda(u-1)\,G(u).$$

The solution of this variable-separable differential equation is : $G(u) = Ce^{\lambda(u-1)t}$. Since the set of events is complete, $G(1) = 1 \Rightarrow C = 1$. Thus

$$G(u) = e^{\lambda(u-1)t} = e^{-\lambda t}.e^{\lambda ut} = e^{-\lambda t}\,\Sigma(\lambda ut)^r/r!, \quad (0 \le r < \infty).$$

$$\therefore \qquad p_k = \text{Coeff. of } u^k/k! = e^{-\lambda t}(\lambda t)^k/k!.$$

5*. Obviously, $P_k \ge 0$, and

$$\sum_{k=0}^{\infty}P_k = \sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\frac{p_j e^{-x_j}(x_j)^k}{k!} = \sum_{j=0}^{\infty}p_j e^{-x_j}\sum_{k=0}^{\infty}\frac{(x_j)^k}{k!} = \sum_{j=0}^{\infty}p_j e^{-x_j}(e^{x_j}) = \sum_{j=0}^{\infty}p_j = 1.$$

It follows that P_k is a p.m.f. of some r.v. Y; in fact $P_k = P(Y=k)$. Now using (1),

$$G(t:Y) = \mathrm{E}(t^Y) = \sum_{k=0}^{\infty}P_k t^k = \sum_{j=0}^{\infty}\sum_{k=0}^{\infty}p_j e^{-x_j}\frac{(tx_j)^k}{k!} = \sum_{j=0}^{\infty}p_j e^{-x_j}\sum_{k=0}^{\infty}\frac{(tx_j)^k}{k!} = \sum_{j=0}^{\infty}p_j e^{-x_j}e^{tx_j} \quad \dots(2)$$

This is the generating function for the distribution (y_i, P_i). Now

$$\mathrm{E}(Y) = G'(1) = [\Sigma p_j e^{-x_j}x_j e^{itx_j}]_{t=1} = \Sigma p_j x_j = \mathrm{E}(X).\ (0 \le j < \infty)$$

$$G''(1) = [\Sigma p_j e^{-x_j}x_j^2 e^{itx_j}]_{t=1} = \Sigma p_j x_j^2 = \mathrm{E}(X^2).\quad (0 \le j < \infty)$$

$\therefore \quad \text{Var}(Y) = G''(1) + G'(1) - [G'(1)]^2 = \text{E}(X^2) + \text{E}(X) - \text{E}^2(X) = \text{Var}(X)\text{E}(X).$

6*.
$$G(t:Z) = G(t:X+Y) = G(t:X).G(t:Y) = [p/(1-qt)][p'/(1-q't)]$$
$$= \frac{pp'}{q-q'}\left\{\frac{1}{1-qt} - \frac{1}{1-q't}\right\}\frac{1}{t} = \frac{pp'}{q-q'}\sum_{r=1}(q^r - q'^r)t^{r-1}$$
$$= \frac{pp'}{q-q'}\sum_{k=0}(q^{k+1} - q'^{k+1})t^k. \quad [r = 1+k]$$
$$P(Z=k) = pp'(q^{k+1} - q'^{k+1})/(q-q'), \qquad q \neq q'$$

Chapter 9 : Characteristic Functions

Sec. 9-12. Page 322

1*.
$$\frac{\Gamma(n+r)}{\Gamma(n)} = \frac{(n+r-1)!}{(n-1)!} = \binom{n+r-1}{r}(r!)$$
$$\therefore \quad \phi(t) = \sum_{r=0}^{\infty}\frac{(it)^r}{r!}\mu'_r = \sum_{r=0}^{\infty}\binom{n+r-1}{r}(it)^r = (1-it)^{-n}.$$

2*. In case c.g.f. exists, it is given by
$$K(t:X) = \sum_{r=1}^{\infty}\frac{(it)^r}{r!}k_r = \sum_{r=1}^{\infty}\frac{(it)^r}{r!}(r-1)!\,n = n\sum_{r=0}^{\infty}\frac{(it)^r}{r} = -n\ln n(1-it) = \ln(1-it)^{-n}.$$

Since $K(t:X) = \ln\phi(t:X)$, it follows that $\phi(t) = (1-it)^{-n}$.

3*. The constant y_0 is determined by normality $\sum f(x) = 1$. Assuming that the given law is p.d.f., we get
$$\mu'_r = \text{E}(X^r) = y_0\int_{-\infty}^{\infty} e^{-|x|}x^r\,dx. \qquad \text{...(1)}$$

If $r = 2n+1$, then the above integrand is an odd function of x, so that $\mu'_{2n+1} = 0$. In particular, $\mu'_1 = \mu = 0$. If $r = 2n$, then the integrand in (1) is an even function of x, and since the central moments are identical with simple moments (as $\mu = 0$) we get
$$\mu_{2n} = 2y_0\int_0^{\infty} e^{-|x|}x^{2n}\,dx = 2y_0\int_0^{\infty} e^{-x}x^{(2n+1)-1}\,dx = 2y_0\,\Gamma(2n+1) = 2y_0\,(2n!).$$

To determine y_0, we use normality $\sum f(x) = 1 = \mu_0$, so putting $n = 0$, we get $1 = 2y_0 \Rightarrow y_0 = 1/2$.

From $\mu_{2n} = 2n\,!$ we get $\sigma^2 = \mu_2 = 2$ so that S.D. $\sigma = \sqrt{2}$.
$$M = \text{E}(|x-0|) = \frac{1}{2}\int_{-\infty}^{\infty}|x|e^{-|x|}\,dx\int_0^{\infty} xe^{-x}\,dx = \Gamma(2) = 1.$$
$$\phi(t) = \text{E}(e^{itx}) = \frac{1}{2}\int_{-\infty}^{\infty} e^{-|x|}e^{itx}\,dx = \frac{1}{2}\int_{-\infty}^{\infty} e^{-|x|}(\cos tx + i\sin tx)\,dx$$
$$= \int_0^{\infty} e^{-x}\cos tx\,dx = \frac{\left[e^{-x}(-\cos tx + t\sin tx)\right]_0^{\infty}}{(1+t^2)} = \frac{1}{1+t^2}.$$

Sec. 9-81. Page 337

1*. $\phi_1(-t) \neq \overline{\phi_1(t)} \Rightarrow \phi_1(t)$ is not a Ch. Function, same is true of $\phi_2(t)$. We may even note that $\phi_2(0) \neq 1$. Since $|\phi_4(t)| = 2 \leq 1, \phi_4(t)$ is not a Ch. Function. Also $\phi_5(0) \to 0$ so that $\phi_5(0) \neq 1$.

$\phi_3(t) = \frac{1}{2}(e^{ibt} + e^{-ibt})$ is the Ch. Function of discrete distribution $f(b) = f(-b) = \frac{1}{2}$.

$\phi_6(t)$ satisfies Bochner's test whence it is a Ch. Function. It is Ch. Function of Lap (1, 0) Dist.

$\phi_7(t)$ is Ch. Function of the degenerate distribution $P(X = 1) = 1,\ P(X \neq 1) = 0$.

2*. $$\phi_X(t) = \frac{e^{it}}{2 - e^{it}} = \frac{1}{2}e^{it}\left(1 - \frac{e^{it}}{2}\right)^{-1} = \frac{1}{2}e^{it}\sum_{r=0}^{\infty}\left(\frac{e^{it}}{2}\right)^r = \sum_{r=0}^{\infty}\left(\frac{1}{2}\right)^{r+1} . e^{i(r+1)t} \qquad \ldots(1)$$

Since $\phi(t) = \text{E}(e^{itX}) = \sum_{r=0}^{\infty} e^{itX} P(X = x)$, we find from (1) that X is discrete with p.m.f.

$$P(X = r + 1) = \left(\frac{1}{2}\right)^{r+1}, r \geq 0, \text{ i.e } P(X = r) = \left(\frac{1}{2}\right)^r, \ r \geq 1.$$

3*. By Inversion Formula, the density f is

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx}\phi(t)\,dt = \frac{1}{2\pi}\int_{-\infty}^{\infty}(\cos tx - i\sin tx)\,e^{-\lambda|t|}\,dt$$

$$= \frac{1}{\pi}\int_0^{\infty}\cos tx \,.\, e^{-\lambda t}\,dt \qquad \text{[Even-odd properties of definite integral]}$$

$$= \frac{1}{\pi}\frac{\left\{e^{-\lambda t}[-\lambda\cos tx + t\sin tx]\right\}_0^{\infty}}{(\lambda^2 + x^2)} = \frac{\lambda}{\pi(\lambda^2 + x^2)}, \quad -\infty < x < \infty.$$

4*. Here $f(x, y, u, v) = f_1(x) f_2(y) f_3(u) f_4(v)$, where $f(t) = \left(\sqrt{2\pi}\right)^{-1}\exp(-t^2/2), -\infty < t < \infty$. It follows that X, Y, U, V are i.i.d. $N(0, 1)$ variates. Firstly let us find Ch. Function of XY.

$$\phi_{XY}(t) = \text{E}(e^{itXY}) = \text{E}\{\text{E}(e^{itXY} \mid y)\} = \text{E}\left\{\text{Ch. Function } f \text{ of } N(0,1) \mid y\right\} = \text{E}\left(e^{-\frac{1}{2}t^2Y^2}\right)$$

$$= \int_{-\infty}^{\infty}\frac{e^{-\frac{1}{2}y^2}}{\sqrt{2\pi}}.e^{-\frac{1}{2}t^2y^2}\,dy = \frac{1}{\sqrt{1+t^2}}\left(\int_{-\infty}^{\infty}\frac{e^{-\frac{1}{2}w^2}}{\sqrt{2\pi}}dw\right) = \frac{1}{\sqrt{1+t^2}}.\left[y\sqrt{1+t^2} = w\right] \quad \ldots(2)$$

$\therefore$ $$\phi(t : XY + UV) = \phi(t : XY)\,\phi(t : UV) = 1/(1 + t^2).$$

This is known to be the Ch. Function of Laplace variate, whose p.d.f. is given by (1).

Note. We can use Inversion theorem to obtain (1) from (2) or we can find the Ch. Function of Z from (1) to obtain (2).

5*. $\phi(t_1, t_2) = E(e^{it_1U+it_2V}) = E[e^{i(t_1X+t_2Y)Z}] = E\{Ee^{i(t_1X+t_2Y)Z} \mid Z = z\}$ [By Double-E Rule]

$$= E\{E[e^{(it_1z)X} . e^{(it_2z)Y} \mid Z = z]\} = E\{E\, e^{(it_1z)X} \mid Z = z)\,(E\, e^{(it_2z)Y} \mid Z = z)\}$$

$$= E\left\{e^{-\frac{1}{2}t_1^2Z^2} . e^{-\frac{1}{2}t_2^2Z^2}\right\} = E\left\{e^{-\frac{1}{2}(t_1^2+t_2^2)Z^2}\right\}$$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-\frac{1}{2}z^2} . e^{-\frac{1}{2}(t_1^2+t_2^2)Z^2}\, dz = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \frac{e^{-\frac{1}{2}w^2}\, dw}{\sqrt{1+t_1^2+t_2^2}} \qquad \left[\sqrt{1+t_1^2+t_2^2}\; z = w\right]$$

$$= \left(1+t_1^2+t_2^2\right)^{1/2} . \left[\text{Since area under } N(0, 1) \text{ curve is unity} : \int_{-\infty}^{\infty} \frac{e^{-\frac{1}{2}w^2}\, dw}{\sqrt{2\pi}} = 1\right]$$

$$\phi(t_1, 0) = \phi_U(t) = (1+t_1^2)^{-1/2}, \phi(0, t_2) = \phi_V(t) = (1+t_2^2)^{-1/2}.$$

Since $\phi(t_1, t_2) \neq \phi(t_1, 0)\,\phi(0, t_2)$, the r.v.s. U and V are not independent.

6*. Here Rect $(\theta) = 1 . I(|\theta| < 1/2) + 0 . I(|\theta| > 1/2)$. By definition of Ch. Function

$$\phi(t_1, t_2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x, y)\, e^{i(t_1x+t_2y)}\, dx\, dy = \frac{1}{2\pi}\int_{-\pi/2}^{\pi/2} e^{it_1x}\, dx \int_{-\pi/2-x}^{\pi/2-x} \left(\frac{e^{i(x+y)} + e^{-i(x+y)}}{2}\right) e^{it_2y}\, dy.$$

Put $y = z - x$, $dy = dz$, then using $e^{i\theta} = \cos\theta + i\sin\theta$ the above reduces to

$$4\pi\phi(t_1, t_2) = \int_{-\pi/2}^{\pi/2} e^{i(t_1-t_2)x} dx . \int_{-\pi/2}^{\pi/2} [e^{i(1+t_2)z} + e^{-i(1-t_2)z}]\, dz$$

$$= 4\int_0^{\pi/2} \cos(t_1 - t_2)\, x\, dx . \int_0^{\pi/2} [\cos(1+t_2)z + \cos(1-t_2)z]\, dz. \text{[by even-odd properties]}$$

$$= \frac{4\left[\sin(t_1-t_2)\,x\right]_0^{\pi/2}}{(t_1 - t_0)}\left\{\frac{\left[\sin(1+t_2)\,z\right]_0^{\pi/2}}{1+t_2} + \frac{\left[\sin(1-t_2)\,z\right]_0^{\pi/2}}{(1-t_2)}\right\}$$

$$= \frac{4\sin\left[(t_1-t_2)\,\pi/2\right]}{t_1 - t_2} . \frac{2\cos(\pi t_2/2)}{(1-t_2^2)}$$

Thus, $\phi(t_1, t_2) = \dfrac{\cos(\pi t_2^2)}{(1-t_2^2)} . \dfrac{\sin[(t_1-t_2)\,\pi/2]}{[(t_1-t_2)\,\pi/2]}$. $\phi_X(t_1) = \dfrac{\sin(\pi t_1/2)}{(\pi t_1/2)}$, $\phi_Y(t_2) = \dfrac{\sin(\pi t_2)}{(1-t_2^2)(\pi t_2)}$.

7*. By Example 2 (§9.81 Worked-out problems) we record that

$$f(x) = (1 - \cos x)/x^2, \quad x \in R; \quad \phi_X(t) = 1 - |t|, \quad |t| \leq 1. \qquad ...(1)$$

Now, consider the discrete distribution

$$P(Y = 0) = \tfrac{1}{2}, P\{Y = (2k-1)\,\pi\} = 2/(2k-1)^2\,\pi^2, \; k = 0, \pm 1, \pm 2, ...$$

We compute Ch. Function of Y

$$\phi_Y(t) = \mathsf{E}(e^{itY}) = \frac{1}{2} + \sum_j p_j e^{itr} = \frac{1}{2} + \sum_k \frac{2}{(2k-1)^2 \pi^2} e^{it(2k-1)\pi}$$

$$= \frac{1}{2} + \frac{4}{\pi^2} \sum_{k=1}^{\infty} \frac{\cos(2k-1)\pi t}{(2k-1)^2} \qquad [e^{i\theta} + e^{-i\theta} = 2\cos\theta] \quad \ldots(2)$$

by considering (0, 1), (–1, 2), (–2, 3), etc.

From Advanced Calculus, recall the Fourier Series for function $|t|$.

$$|t| = \frac{1}{2} - \frac{4}{\pi^2} \sum_{k=1}^{\infty} \frac{\cos(2k-1)\pi t}{(2k-1)^2}, \quad -1 \le t \le 1. \qquad \ldots(3)$$

From (2) and (3) $\phi_Y(t) = 1 - |t|, \ |t| \le 1.$

We conclude : $f_X(x) \ne f_Y(x)$, but $\phi_X(t) = \phi_Y(t)$, for $|t| \le 1$.

8*. Recall from Trigonometry : $\sum_{r=0}^{k-1} e^{[2\pi in/k]^r} = \begin{cases} k, & \text{if } k \mid n \\ 0, & \text{otherwise} \end{cases}$ ...(1)

Now, $$\phi(t) = \mathsf{E}(e^{itX}) = \sum_{n=0}^{\infty} p_n e^{itn} \Rightarrow \phi\left(\frac{2\pi r}{k}\right) = \sum_{n=0}^{\infty} e^{in(2\pi r/k)} p_n$$

$$\sum_{r=0}^{k-1} \phi\left(\frac{2\pi r}{k}\right) = \sum_{n=0}^{\infty} p_n \sum_{r=0}^{k-1} e^{2\pi inr/k} = k \sum_{n=0}^{\infty} P(X = n;\, k \mid n) = kP\{X \equiv 0 \bmod (k)\}$$

This is equivalent to the result stated.

Chapter 10 : Some Probability Inequalities

Sec. 10-21. Page 349

1*. If $Z = (X - \mu)/\sigma$ is standard r.v. then Chebyshev's inequality is

$P\{|Z| \ge k\} \le 1/k^2$, or $P\{|Z| < k\} > 1 - (1/k^2)$. Here $Z = (X - 50)/10$. ...(1)

(a) $P\{|Z| \ge 3/2\} \le 4/9$ (upper bound). (b) $P(Z \ge 3/2) + P(<-3/2) = P(|Z| \ge 3/2) \le 4/9$.

(c) $P(|Z| < 2) > 1 - \frac{1}{4} = \frac{3}{4}$. (d) $P(-2 < Z < 2) = P(|Z| < 2) = \frac{3}{4}$.

(e) $P(|Z| \ge t/10) \le 100/t^2 \Rightarrow (100/t^2) \ge (100)^2$, i.e. $t \ge 100$.

2*. By Chebyshev's Inequality : $P\{|Z_k| \le t\sqrt{n}\} \ge 1 - (1/nt^2)$ ...(1)

Let $A_k = \{|Z_k| \le t\sqrt{n}\}$ Recall Bon-ferroni Inequality (§1-55(5))

$$P(A_1, A_2, \ldots, A_n) \ge \sum P(A_k) - (n-1) \qquad \ldots(2)$$

As per (1), $P(A_k) \ge 1 - (1/nt^2)$, using it in (2),

$$P\{|Z_k| \le \sqrt{n}\,t, 1 \le k \le n\} \ge n[1 - (1/nt^2)] - n + 1 = 1 - (1/t^2)$$

3*. Let $\mathsf{E}(X_k) = \mu$, Var $(X_k) = \sigma^2$; then

$$\mathsf{E}(Z_n) = \frac{6}{n(n+1)(2n+1)} \sum_{k=1}^{n} k^2\, \mathsf{E}(X_k) = \frac{6\mu}{n(n+1)(2n+1)} \cdot \frac{n(n+1)(2n+1)}{6} = \mu$$

$$\text{Var}(Z_n) = \left[\frac{6}{n(n+1)(2n+1)}\right]^2 \sum_{k=1}^{n} k^4 \text{Var}(X_k)$$

$$= \frac{36\sigma^2}{[n(n+1)(2n+1)]^2} \frac{n(n+1)(2n+1)(3n^2+3n-1)}{30} = \frac{6\sigma^2}{5}\frac{3n^2+3n-1}{n(n+1)(2n+1)}$$

Now $P\{|Z_n - E(Z_n)| > \varepsilon\} \le [\text{Var}(Z_n)/\varepsilon^2]$. [Chebyshev's inequality] ...(1)

Since $\text{Var}(Z_n) \to 0$, as $n \to \infty$, (1) reduces to $P\{|Z_n - \mu| > \varepsilon\} \to 0$, i.e. $Z_n \xrightarrow{P} E(X_1) = \mu$.

4*. $\text{Var}(X) = E(X^2) - E^2(X) = 4 \Rightarrow \sigma = 2.$

$$p = P(-2 < X < 8) = P(-5 < X-3 < 5) = P(|X-3| < 5).$$

Chebyshev's inequality : $P(|X-\mu| < c) \ge 1 - (\sigma^2/c^2)$ provides $p \ge 1 - (4/25)$, i.e. $p \ge 21/25$.

5*. Here $E(\bar{X}) = \mu = 5,\ \text{Var}(\bar{X}) = \sigma^2/n = 1/n,\ (c = 10^{-3})$

$$P\{|X-\mu| < c\} \ge 1 - (\sigma^2/c^2) \Rightarrow P\{|\bar{X} - 5| < 0.001\} \ge 1 - (10)^6/n.$$

6*. $E(X) = (-a)\frac{1}{8} + (a).\frac{1}{8} + 0.\frac{3}{4} = 0, E(X^2) = a^2.\left(\frac{1}{8}\right) + a^2\left(\frac{1}{8}\right) = \frac{1}{4}a^2\ \mu = 0,$

$$\sigma^2 = \text{Var}(X) = E(X^2) = \frac{a^2}{4},\ \text{i.e. } \sigma = \frac{a}{2}.$$ Now by Chebyshev's inequality.

$$P(|X| \ge 2\sigma) = P(|X-\mu| \ge a) \le 1/4.$$

Actually, $p = P\{X \ge 2\sigma\} = 1 - P\{|X| < a\} = 1 - P(-a < X < a) = 1 - f(0) = 1 - 3/4 = 1/4.$

Hence Chebyshev's upper bound coincides with actual value and so the upper bound is attained.

7*. Here $E(X) = \frac{1}{8}(-1) + \left(\frac{6}{8}\right)0 + (1)\frac{1}{8} = 0,\ E(X^2) = \frac{1}{8} + \frac{1}{8} = \frac{1}{4}.\ \text{Var}(X) = E(X^2) = \frac{1}{4}.$

By Chebyshev's inequality : $p = P\{|X-\mu| \ge 2\sigma\} \le \sigma^2/4\sigma^2 = 1/4.$

Exact prob. $p = P(|X| \ge 1) = 1 - P(|X| < 1) = 1 - P(-1 < X < 1) = 1 - (6/8) = 1/4.$

The two result coincide; thus in general, Chebyshev's inequality cannot be improved.

8*. $p = P\{\mu_X - 2\sigma \le X \le \mu_X + 2\sigma\} = P[|X-\mu| \le 2\sigma] > 1 - (\sigma^2/4\sigma^2) = 1 - \frac{1}{4} = 0.75$ ...(1)

Thus $p > 0.75 \Rightarrow p \ne 0.60$. So there does not exist any variate X for which (1) holds.

9*. Here C.V. = σ/μ. Hence by Chebyshev's inequality

$$P\{|\bar{X} - \mu| \le 0.05\mu\} \ge 1 - \frac{\sigma^2/n}{(0.05\mu)^2} = 1 - \frac{400}{n}.\frac{\sigma^2}{\mu^2} = 1 - \frac{4}{n}.$$

By hypothesis, $1 - (4/n) \ge 0.90 \Rightarrow n \ge 40$.

10*. $P(X_k = 1) = 1/(k+1) = E(X_k),\ k = 1, 2, 3, \ldots$

$$\text{Var}(X_k) = E(X_k^2) - E^2(X_k) = [1/(k+1)] - 1/(k+1)^2 = k/(k+1)^2$$

$$E(S_n) = \Sigma E(X_k) = (1/2) + (1/3) + \ldots + (1/n)$$

$$\text{Var}(S_n) = \Sigma \text{Var}(X_k) = \Sigma\{[1/(1+k)] - [1/(1+k)^2]\}.\ 1 \le k \le n.$$

Recall : $\lim\limits_{n\to\infty} [1 + 1/2 + 1/3 + ... + (1/n) - \ell n\, n] = \gamma = 0.58.$ [Euler's constant] ...(2)

$\therefore \quad \lim\{[1 + (\frac{1}{2}) + (\frac{1}{2}) + ... + (\frac{1}{n})](\ell n\, n)^{-1}\} = 1.$ [Divide (2) by $\ell n\, n$ and let $n \to \infty$] ...(3)

Note that $\Sigma(1/k)$ is divergent and $\Sigma(1/k^2)$ is convergent. Now let,

$$Z_n = (S_n / \ell n\, n),\ \theta_n = E(Z_n) = E(S_n)/(\ell n\, n),\ \text{Var}(Z_n) = \text{Var}(S_n)/(\ell n\, n)^2.$$

Obviously Var $(S_n) \to 0$ as $n \to \infty$. Since $\lim \theta_n = 1$, so $|\theta_n - 1| < 1/2\ \varepsilon$ for all $n \ge n_0$.

By Chebyshev's inequality : $P\{|Z_n - \theta_n| < \frac{1}{2}\varepsilon\} \ge 1 - 4[\text{Var}(Z_n)/\varepsilon^2]$ ...(4)

Now $|Z_n - 1| = |(Z_n - \theta_n) + (\theta_n - 1)| \le (|Z_n - \theta_n| = |\theta_n - 1|)$ [Triangle inequality]

Since $|\theta_n - 1| < \frac{1}{2}\varepsilon$, for all $n \ge n_0$, hence when $n \ge n_0$, $|Z_n - \theta_n| < \frac{1}{2}\varepsilon \Rightarrow |Z_n - 1| < \varepsilon,\ \forall\ n \ge n_0$.

i.e. $\quad P\{|Z_n - \theta_n| < \frac{1}{2}\varepsilon\} \le P\{|Z_n - 1| < \varepsilon\}$

or $\quad P\{|Z_n - 1| < \varepsilon\} \ge 1 - 4\,\text{Var}(Z_n)/\varepsilon^2,\ \forall\ n \ge n_0,$ [by (4)] ...(5)

Thus, as $n \to \infty$, and hence eventually $n \ge n_0$, the R.H.S. of (5) tends to unity. This establishes (1).

Sec. 10-41. Page 356

1*. $P\{X \ge t\} = P\{(X+c) \ge (t+c)\} \le P\{X+c)^2 \ge (t+c)^2\} \le E[(X+c)^2]/(c+t)^2 = \phi(c)$ ...(1)

where we used C-β inequality §10-30 (1) with $g(X) = (X + c)^2$.

Now if $\mu = 0$, $E(X+c)^2 = E(X^2 + 2cX + c^2) = \sigma^2 + c^2$; so we let

$$\phi(c) = \frac{\sigma^2 + c^2}{(c+t)^2},\ \phi'(c) = \frac{2(ct - \sigma^2)}{(c+t)^3},\ \phi''(c) = \frac{2(t^2 - 2ct + 3\sigma^2)}{(t+c)^4}.$$

$$\phi'(c) = 0 \Rightarrow c = \frac{\sigma^2}{t},\ \phi''\left(\frac{\sigma^2}{t}\right) = \frac{(t^2 + \sigma^2)\,t^4}{(t^2 + \sigma^2)^4} > 0.$$

Thus, $\phi(c)$ is minimum at $c = \sigma^2/t$, and min $\phi(c) = \sigma^2/(\sigma^2 + t^2)$.

Now, result (1) is true $\forall\ c$; it particular it is true for $c = \sigma^2/t$.

$$\therefore \quad P(X \ge t) \le \sigma^2/(\sigma^2 + t^2).$$

2*. Let us write Cov $(X_i, X_j) = \sigma_{ij}$. Now

$\text{Var}(\Sigma p_i X_i) = \sum\limits_i\sum\limits_j \text{Cov}(p_i X_i, p_j X_j) = \sum\limits_i\sum\limits_j p_i p_j \sigma_{ij} < \sum\limits_i\sum\limits_j p_i p_j \sigma_i \sigma_j$ [§10-14] ...(1)

Now, $\sum\limits_i\sum\limits_j p_i p_j \sigma_i \sigma_j = \left(\sum\limits_i p_j \sigma_i\right)\left(\sum\limits_i p_j \sigma_j\right) = (\Sigma p_i \sigma_i)^2$ ...(2)

From $(\sigma_i - \sigma_j)^2 \geq 0$ we get $\sigma_i \sigma_j \leq \frac{1}{2}(\sigma_i^2 + \sigma_j^2)$ and thus

$$\sum_i \sum_j p_i p_j \sigma_i \sigma_j \leq \frac{1}{2} \sum_i \sum_j p_i p_j (\sigma_i^2 + \sigma_j^2) = \frac{1}{2}\left[\sum_i p_i \sum_j p_j \sigma_j^2 + \sum_j p_j \sum_j p_i \sigma_i^2\right] = \sum p_i \sigma_i^2 \quad \ldots(3)$$

Combining (1), (2) and (3) we get : $\sigma^2(\Sigma p_i X_i) \leq (\Sigma p_i \sigma_i)^2 \leq \Sigma p_i \sigma_i^2$.

Taking positive square roots, we obtain : $\sigma(\Sigma p_i X_i) \leq \Sigma(p_i \sigma_i) \leq (\Sigma p_i \sigma_i^2)^{1/2}$.

Note. We have assumed that all the series involved are absolutely convergent.

3*. Let Cov $(X_i, X_j) = \sigma_{ij}$. Note that $(\sigma_i - \sigma_j)^2 \geq 0$ gives $\sigma_i \sigma_j < \frac{1}{2}(\sigma_i^2 + \sigma_j^2)$. ...(1)

$$\text{Var}(S_n) = \text{Cov}(S_n, S_n) = \text{Cov}(\Sigma X_i, \Sigma X_j) = \Sigma\Sigma \text{Cov}(X_i, X_j) = \Sigma\Sigma\sigma_{ij}$$

$$\leq \Sigma\Sigma\sigma_i\sigma_j < \tfrac{1}{2}\Sigma\Sigma(\sigma_i^2 + \sigma_j^2) = n\,\Sigma\sigma_i^2 \quad \text{[By Schwarz Inequality and by (1)]}$$

Thus, $\quad \text{Var}(S_n/n) = n^{-2}\,\text{Var}(S_n) \leq n^{-1}(\Sigma\sigma_i^2), \; 1 \leq i \leq n.$

Remark. You may even use *Mathematical Induction* to prove this result.

$$\text{Cov}(S_k, X_{k+1}) = \Sigma_i \text{Cov}(X_i, X_{k+1}) \leq \Sigma\sigma_i\sigma_{k+1} \leq \tfrac{1}{2}\Sigma(\sigma_i^2 + \sigma_{k+1}^2) = \tfrac{1}{2}[(\Sigma_i \sigma_i^2) + k\sigma_{k+1}^2].$$

4*. To simplify writing, we put $Y = |X|$. Since expectation of a non-negative quantity is necessarily non-negative, we must have

$$0 \leq E\{Y^a(\lambda Y^b + 1)^2\} = \lambda^2 E(Y^{a+2b}) + 2\lambda E(Y^{a+b}) + E(Y^a) \text{ i.e. } \lambda^2 V_{a+2b} + 2\lambda V_{a+b} + V_a \geq 0. \; [V_r = E(Y^r)]$$

Since this Quadratic in λ is non-negative, its discriminant is non-positive. Consequently

$$(V_{a+b})^2 \leq V_a \cdot V_{a+2b}$$

Let $a = m$, $a + 2b = n$, $[a + b = (m + n)/2]$ then this result provides the stated result (1).

Convex Function. $g(\theta)$ is called *convex function* of θ if, for every pair (α, β).

$$g\left[\frac{1}{2}(\alpha + \beta)\right] \leq \left[\frac{1}{2} g(\alpha) + g(\beta)\right].$$

In the present case, using (1) :

$$g\left[\tfrac{1}{2}(m+n)\right] = \ln E[Y^{(m+n)/2}] \leq \ln\,[E(Y^m)E(Y^n)]^{1/2} \leq \tfrac{1}{2}\{\ln E(Y^m) + \ln E(Y^n)] = \tfrac{1}{2}\{2g(m) + g(n)\}.$$

It follows that $\ln(|X|^n)$ is convex.

Chapter 11 : Convergence. Limiting MGF. Laws of Large Numbers

Sec. 11-31. Page 370

1*. $M_n(t) = (q + pe^t)^n = [1 + p(e^t - 1)]^n = \left[1 + \dfrac{\lambda(e^t - 1)}{n}\right]^n$

$$\lim_{n\to\infty} M_n(t) = \lim_{n\to\infty}\left[1 + \frac{\lambda(e^t - 1)}{n}\right]^n = e^{\lambda(e^t - 1)}. \quad \text{[m.g.f. of Pois}(\lambda)\text{ \& by Euler's limit]}$$

By Continuity theorem of m.g.f., $X \to$ Pois (λ) under stated conditions.

2*. We work with standardized variate; $X^* = (X-\lambda)/\sqrt{\lambda}$. Now

$$\begin{aligned} M(t:X^*) &= M\left(t:\frac{X}{\sqrt{\lambda}}-\sqrt{\lambda}\right) = e^{-\sqrt{\lambda}t} M\left(\frac{t}{\sqrt{\lambda}}:X\right) = e^{-\sqrt{\lambda}t}\exp[\lambda(e^{t/\sqrt{\lambda}}-1)] \\ &= \exp\{-\sqrt{\lambda}t + \lambda(e^{t/\sqrt{\lambda}}-1)\} \\ &= \exp\{-\sqrt{\lambda}t + \lambda(t/\sqrt{\lambda} + t^2/2\lambda + t^3/3!\lambda^{3/2} + \ldots)\} \\ &= \exp\{t^2/2 + 0(\lambda^{1/2})\} \end{aligned}$$

Thus, $\quad M(t:X^*) \to e^{t^2/2}$. [m.g.f. of $N(0, 1)$] as $\lambda \to \infty$.

3*. $\quad M(t:X_n) = (1-t/\lambda)^{-n}$, [§ 8-16(12)].

$\therefore \quad M(t:X_n/n) = M(t/n:X_n) = M(1-t/\lambda n)^{-n}$

$$\lim_{n\to\infty} M(t:Y) = \lim_{n\to\infty}(1-t/\lambda n)^{-n} = e^{t/\lambda},$$ [mgf of degenerate r.v.], by Euler's limit.

Thus, the limiting distribution of Y is $P\{Y = 1/\lambda\} = 1$, $P\{Y \neq 1/\lambda\} = 0$.

4*. Since $M(t:X) = [pe^t/(1-qe^t)]^k$, by § 8-16(7), it follows that

$$M(t:2pX) = M(2pt:X) = \{pe^{2pt}/(1-qe^{2pt})\}^k.$$

Since $1 - qe^{2pt} = 1 - q(1 + 2pt + 2p^2t^2 + \ldots) = p - pq(2t + 2pt^2 + \ldots)$

$\therefore \quad M(t:2pX) = \{e^{2pt}/[1-q(2t+2pt^2+\ldots)]\}^k \to \{1/(1-2t)\}^k$ as $p \to 0, q \to 1$.

Since $(1-2t)^{-k}$ is the m.g.f. of χ^2_{2k} [§8-16(14)], it follows that $2pX \to \chi^2_{2k}$ as $p \to 0$.

Sec. 11-62. Page 377

1*. Here $E(X_j) = \mu$, Var $(X_j) = \sigma^2$. Let $T_n = X_1^2 + X_2^2 + \ldots + X_n^2$. So $E(T_n) = nE(X_1^2) = n[\sigma^2 + \mu^2]$.

In a way, first moment of X^2 exists, hence by Khinchin's WLLN;

$$(T_n/n) = (X_1^2 + X_2^2 + \ldots + X_n^2)/n \xrightarrow{p} E(X_1^2) = \sigma^2 + \mu^2; \Rightarrow c = \mu^2 + \sigma^2.$$

2*. $E(X_k) = \sum_{r=1}^{\infty} 2^{r-2\ln r}2^{-r} = \sum_{r=1}^{\infty}\left(\frac{1}{4}\right)^{\ln r} = \sum_{r=1}\left(\frac{1}{r}\right)^{\ln 4} \qquad [\because a^{\ln n} = (n)^{\ln a}]$

Since $\ln 4 = 1.39 > 1$, $E(X_k)$ is finite [Convergent hyperharmonic series $(1/n)^p$, $p > 1$]. Since $\langle X_k \rangle$ are i.i.d. variates, with $E(X_k) < \infty$, the WLLN holds for this sequence, by Khintchin's Theorem.

3*. Here $E(X_{nr}) = \sum_{k=1}^{\infty}\frac{(r+1)^{k/2}}{r}\cdot\frac{r}{(1+r)^k} = \sum_{k=1}^{\infty}\rho^k = \frac{\rho}{1-\rho} = \frac{1}{\sqrt{1+r}-1}. \quad \left[\rho = \frac{1}{\sqrt{1+r}}\right]$

$$E(X_{nr}^2) = \sum_{k=1}^{\infty}\frac{(r+1)^k}{r^2}\cdot\frac{r}{(1+r)^k} = \sum_{1}^{\infty}\left(\frac{1}{r}\right) \to \infty.$$

We conclude that $\mathsf{E}(X_n)$ exists for positive r but Var (X_n) does not exist. However, infinite variance need not rule out the applicability of WLLN. We try Markov's Rule :

$$\mathsf{E}\left(|X_{nr}|^{1+\delta}\right) = \sum_{k=1}^{\infty}\left|\frac{(r+1)^{k/2}}{r}\right|^{1+\delta}\frac{r}{(1+r)^k} = \frac{1}{r^{\delta}}\sum_{k=1}^{\infty}\frac{1}{\rho^{(1-\delta)k}} < \infty, \qquad [0<\delta<1].$$

Thus, the sequence $\langle X_{nr}\rangle$ admits of WLLN by Markov's rule.

4*. Let $Y = -\ln X$, where $X \sim U(0, 1)$; then

$$F_Y(y) = P(Y \le y) = P(X \ge e^{-y}) = \int_{e^{-y}}^{1} 1\,dx = 1 - e^{-y}. \quad f_Y(y) = F'(y) = e^{-y}, \quad y > 0.$$

$\mathsf{E}(Y) = \text{Var}(Y) = 1$. The sequence $\langle Y_k\rangle$, $k = 1, 2, \ldots$, is such that Y_k are i.i.d with finite mean $\mathsf{E}(Y_i) = 1$. Hence by Khintchin's WLLN,

$$\frac{(\Sigma Y_k)}{n} \xrightarrow{p} \mathsf{E}(Y_1) \cdot \left[\ln G_n = \frac{(\Sigma \ln X_k)}{n} = \frac{-(\Sigma Y_k)}{n}\right]$$

or $\quad -\ln G_n \xrightarrow{p} 1$, i.e $G_n \xrightarrow{p} e^{-1}$. This yields $c = 1/e$.

5*. (*i*) $\mathsf{E}(\overline{X}_n) = \dfrac{1}{n}\mathsf{E}(X_1 + \ldots + X_n) = \dfrac{1}{n}\cdot\displaystyle\sum_{j=1}^{n}\left(\frac{0+j}{2}\right) = \frac{1}{2n}(1+2+\ldots+n) = \frac{n+1}{4}.$

As $n \to \infty$, $\mathsf{E}(\overline{X}_n)$ increases indefinitely and there is no stability of the arithmetic mean. Further

$$\text{Var } X_j = \frac{(j-0)^2}{12}; \quad B_n = \Sigma\sigma_j^2 = \Sigma\left(\frac{j^2}{12}\right) = \frac{n(n+1)(2n+1)}{72}.$$

Obviously, $(B_n/n^2) \not\to 0$, hence $\langle X_N\rangle$ does not obey the WLLN.

(ii) $\quad \mathsf{E}(\overline{Y}_n) = \dfrac{\mathsf{E}(Y_1 + \ldots + Y_n)}{n} = 0$, since $\mathsf{E}(Y_j) = \dfrac{1}{2}[j + (-j)] = 0.$

Var $Y_j = [j - (-j)]^2/12 = j^2/3$. It follows that variance of the variates Y_j are not limited by some fixed constant but increase indefinitely with an increase in n. It follows that $\overline{Y}_n$ will not converge to zero in probability ; i.e. $\langle Y_n\rangle$ does not obey the WLLN.

6*. $\mathsf{E}(X_k) = \frac{1}{2}(2k-1)^{1/2} - \frac{1}{2}(2k-1)^{1/2} = 0$. $\text{Var}(X_k) = \mathsf{E}(X_k)^2 = \frac{1}{2}(2k-1) + \frac{1}{2}(2k-1) = 2k-1.$

$$B_n = \sum_{k=1}^{n}\text{Var}(X_k) = \sum_{k=1}^{n}(2k-1) = n^2$$

Since $\ln(B_n/n^2) \not\to 0$, it follows that for $\langle X_n\rangle\}$, the WLLN does not hold.

7*. Here $\mathsf{E}(X_k) = 3^k.3^{-(2k+2)} - 3^k.3^{-(2k+2)} = 0, \ 1 \le k \le n.$

$$\mathsf{E}(X_k^2) = 3^{2k}.3^{-2k-2} + 3^{2k}.3^{-2k-2} = 2/9 = \text{Var}(X_k). \quad B_n = \Sigma\,\text{Var}(X_k) = (2/9)n.$$

As $(B_n/n^2) = (2/9n) \to 0$ as $n \to \infty$, it follows that WLLN holds for the sequence $\langle X_n\rangle$.

8*. Here $\mathsf{E}(X_n) = \frac{1}{2}(1-2^{-n}) - \frac{1}{2}(1-2^{-n}) - \frac{1}{2}(1-2^{-n}) + 2^{-n}.2^{-n-1} - 2^{-n}(2^{-n-1}) = 0.$

$E(X_n^2) = \frac{1}{2}(1-2^{-n}) + \frac{1}{2}(1-2^{-n}) + 2^{-2n}.2^{-n-1} + 2^{-2n}.2^{-n-1} = 1-2^{-n}+2^{-3n} = \text{Var}(X_n).$

$B_n = \Sigma\sigma_r^2 = \Sigma(1-2^{-r}+2^{-3r}) = n-(1-2^{-n})+(1/7)(1-8^{-n}).\ (1 \le r \le n).$ [Use G.P.]

$$\frac{B_n}{n^2} = \frac{1}{n} + \frac{1}{2^n n^2} - \frac{6/7}{n^2} - \frac{1}{7}\frac{1}{8^n}\frac{1}{n^2} \to 0, \text{ as } n \to \infty.$$

It follows that the WLLN holds for the given sequence.

9*. (*i*) We show that $E(X^2) < A^2 \Rightarrow E(X) < \infty$. Assuming X to be continuous with density $f(x)$, we have using C-S in equality.

$$E\,|\,(X_k)\,|^2 = \left|\int_{-\infty}^{\infty}\left[x\sqrt{f(x)}\right]\left[\sqrt{f(x)}\right]dx\right| \le \int_{-\infty}^{\infty} x^2 f(x)\,dx \,.\int_{-\infty}^{\infty} f(x)\,dx = E(X^2).1 < A^2.$$

Thus $E(X_k) < A$, $\forall\ k$. It follows that the sequence $\langle\sigma_k^2\rangle$ is uniformly bounded, because

$$\sigma_k^2 = E(X_k^2) - E^2(X_k) \le E(X_k^2) + E^2(X_k) \le 2A^2.$$

Now $\quad \text{Var}(S_n) \quad = \sum_{k=1}^{n}\sigma_k^2 + 2\sum_{i=1}^{n-1}\sigma_i.\sigma_{i+1}\,\rho_{i,i+1} \qquad [\sigma_{ij} = \sigma_i\,\sigma_i\,\sigma_{ij}]$

$$\le 2nA^2 + 4A^2\sum_{i=1}^{n-1}\rho_{i,i+1} \le 2n\,A^2 + 4A^2(n-1).$$

$\therefore \qquad \text{Var}\,\dfrac{(S_n)}{n^2} \le \dfrac{(6n-4)A^2}{n^2} \to 0$ as $n \to \infty$. Thus $\langle X_n\rangle$ obeys WLLN.

(ii) $\quad \text{Var}\left[\frac{1}{2}(X_k + X_{k+1})\right] = \frac{1}{4}\left[\sigma_k^2 + \sigma_{k+1}^2 + 2\rho\,\sigma_k\,\sigma_{k+1}\right] \le \frac{1}{4}(\sigma_k + \sigma_{k+1})^2$

$\therefore \quad P\left\{\left|\frac{1}{2}(X_k + X_{k+1}) - \frac{1}{2}E(X_k + X_{k+1})\right| \ge \sigma_k + \sigma_{k+1}\right\} \le \dfrac{\text{Var}\frac{1}{2}(X_k + X_{k+1})}{(\sigma_k + \sigma_{k+1})^2} \le \frac{1}{4}.$

10*. Recall : $\text{Var}(aX + bY) = a^2\sigma_x^2 + b^2\sigma_y^2 + 2ab\,\sigma_{xy} \le a^2\,\sigma_x^2 + b^2\,\sigma_y^2$, if $\sigma_{xy} < 0$.

$$\therefore \quad 0 \le \text{Var}\left(\frac{X_1 + X_2 + \ldots + X_n}{n}\right) < \frac{1}{n^2}\sum_{i=1}^{n}\text{Var}(X_i) < \frac{A}{n}. \qquad \ldots(1)$$

where A is the upper bound of Var (X_i), for all $i = 1, 2, \ldots, n$.

The relation : lim var $[(S_n)/n]$ = lim $(B_n/n^2) = 0$, as $n \to \infty$ follows from (1). Thus, the WLLN is applicable to this sequence.

11*. Here $T_n = S_1 + S_2 + \ldots + S_n$ so

$$T_n = X_1 + (X_1 + X_2) + \ldots + (X_1 + X_2 + \ldots + X_n) = nX_1 + (n-1)X_2 + \ldots + 2X_{n-1} + X_n.$$

$$\therefore \quad B_n = \text{Var}(T_n) = n^2\sigma^2 + (n-1)^2\sigma^2 + (n-2)^2\sigma^2 + \ldots + 2^2\sigma^2 + \sigma^2 = \sigma^2(1^2 + 2^2 + \ldots + n^2)$$

$$= n(n+1)(2n+1)\sigma^2/6.$$

Since $(B_n/n^2) \to \infty$, i.e. $(B_n/n^2) \not\to \infty$ as $n \to \infty$, it follows that the necessary condition for WLLN to hold is not satisfied. Thus, WLLN does not hold for $\langle S_n\rangle$.

For the sequence $\{a_n S_n\}$, we have

$$T_n = a_1 S_1 + a_2 S_2 + \dots + a_n S_n = a_1 X_1 + a_2 (X_1 + X_2) + \dots + a_n (X_1 + X_2 + \dots + X_n)$$
$$= (a_1 + a_2 + \dots + a_n) X_1 + (a_2 + a_3 + \dots + a_n) X_2 + \dots + (a_{n-1} + a_n) X_{n-1} + a_n X_n$$

$$B_n = \text{Var}(T_n) = (a_1 + \dots + a_n)^2 \sigma^2 + (a_2 + \dots + a_n)^2 \sigma^2 + \dots + (a_{n-1} + a_n)^2 \sigma^2 + a_n^2 \sigma^2$$
$$= \left\{(a_1^2 + 2a_2^2 + \dots + na_n^2) + 2(a_1 a_2 + \dots + a_1 a_n) + 4(a_2 a_3 + \dots + a_2 a_n) + \dots + 2(n-1) a_{n-1} a_n\right\} \sigma^2$$

$$\therefore (B_n / n^2) = \left\{(\Sigma i a_i^2) + 2(a_1 a_2 + \dots + a_1 a_n) + 4(a_2 a_3 + \dots + a_2 a_n) + \dots + 2(n-1) a_{n-1} a_n\right\} (\sigma^2 / n^2) \quad \dots(1)$$

Because $na_n \to 0$ (given hypothesis), hence by Cauchy Theorem on limits

$$\lim \left[\frac{(a_1 + 2a_2 + \dots + na_n)}{n}\right] = 0$$

Again $na_n \to 0 \Rightarrow a_n = \varepsilon / n$, so that $a_n \to 0$ as $n \to \infty$; whence

$$\lim \left[\frac{(a_1 + a_2 + \dots + a_n)}{n}\right] = 0. \qquad \text{[by Cauchy Limit Theorem]}$$

Let $\theta_n = (a_1 + a_2 + \dots + a_n)(a_1 + 2a_2 + \dots + na_n) / n^2$

$$= [\Sigma (i a_i^2) + 3a_1 a_2 + 4a_1 a_3 + \dots + (n+1) a_1 a_n + \dots + (2n-1) a_{n-1} a_n] / n^2.$$

As $n \to \infty$, $\theta_n \to 0$ (by two preceding limits), hence in (1), $(B_n / n^2) < \theta_n \sigma^2 \to 0$ as $n \to \infty$. Thus, WLLN holds for the sequence $\{a_n S_n\}$ provided $na_n \to 0$.

Chapter 12 : *Bernoulli Binomial Distribution*

Sec. 12-12. Page 386

1*. Let the variates X and Y denote the number of boys and girls in the family ; then

(i) $P\{X \geq 1\} = 1 - P(X = 0) = 1 - \binom{4}{0}\left(\frac{1}{2}\right)^4 = 1 - \frac{1}{16} = \frac{15}{16}.$

(ii) $P\{(X \geq 1) \cap (Y \geq 1)\} = 1 - P\{(X \geq 1) \cap (Y \geq 1)\}^c = 1 - P\{(X = 0) \uplus (Y = 0)\}$

$$= 1 - \{P(X = 0) + P(Y = 0)\} = 1 - \left(\frac{1}{16} + \frac{1}{16}\right) = \frac{7}{8}.$$

2*. Let G and B denote the birth of a girl and a boy, then $P(G) = 1/2 = p = P(B)$. The number X of girls in the family with 5 children is bin $(5, p)$. Hence

$$P(X = k) = \binom{5}{k} q^{5-k} p^k = \frac{5!}{k!(5-k)!} \cdot \frac{1}{32}, \quad k = 0, 1, 2, 3, 4, 5.$$

Let L_g and L_b denote the events : at least one child is a girl and at least one child is a boy. Then since X denotes the number of girls and $5 - X$ denotes the number of boys, so

$$P\{L_g \mid L_b\} = P\{X \geq 1\} \mid \{(5 - X) \geq 1\} = P\{X \geq 1 \mid X \leq 4\} = P\{1 \leq X \leq 4\} / P\{X \leq 4\} \quad \dots(1)$$

$$P\{1 \leq X \leq 4\} = p_1 + p_2 + p_3 + p_4 = \frac{5}{32} + \frac{10}{32} + \frac{10}{32} + \frac{5}{32} = \frac{30}{32} \quad [p_r = P(X = r)] \quad \dots(2)$$

$$P\{X \le 4\} = 1 - P(X > 4) = 1 - P(X = 5) = 1 - (1/32) = 31/32 \qquad \ldots(3)$$

$\therefore \quad P\{L_g \mid L_b\} = 30/31.$ [by (1), (2) and (3)]

3*. $p = P$ {target is detected in a single scan}, $= 0.1$. Now

(a) $P(A) = \sum_{2}^{4} \binom{4}{r} p^r q^{4-r} = 6p^2 q^2 + 4p^3 q + p^4 = 6\left(\frac{1}{10}\right)^2 \left(\frac{9}{10}\right)^2 + 4\left(\frac{1}{10}\right)^3 \left(\frac{9}{10}\right) + \left(\frac{1}{10}\right)^4$

$= 522/10^4 = 0.0522.$

(b) $P(B) = 1 - P$ {zero detection} $= 1 - q^{20} = 1 - (9/10)^{20} = 1 - 0.121 = 0.879.$

4*. Here $p = (6/10)$ (prob. of a good bulb), $q = (4/10)$. Let $L_1 =$ {At least one bulb is good} Now room shall have light if at least one bulb is good. Hence

$$P\{L_1\} = P(X \ge 1) = P(X = 1) + P(X = 2) + P(X = 3)$$

$$= \binom{3}{1}\left(\frac{6}{10}\right)\left(\frac{4}{10}\right)^2 + \binom{3}{2}\left(\frac{6}{10}\right)^2\left(\frac{4}{10}\right) + \binom{3}{3}\left(\frac{6}{10}\right)^2$$

$$= (288 + 432 + 216)/1000 = 936/1000 = 0.936.$$

5*. Here $p = P(A) = 3/5$, $q = P(B) = 2/5$, $n = 5$. The Probability that A wins r games is

$$P(X = r) = \binom{5}{r}\left(\frac{3}{5}\right)^r \left(\frac{2}{5}\right)^{5-r}.$$

If $L_3 = \{A$ wins at least 3 games out of 5$\} = \{X \ge 3\}$, then

$$P(L_3) = \sum_{r=3}^{5} P(X = r) = \sum_{r=3}^{5} \binom{5}{r} \frac{3^r . 2^{5-r}}{5^5} = \frac{3^3}{5^5}(40 + 30 + 9) = \frac{27 \times 79}{3125} = \frac{2133}{3125} = 0.68.$$

6*. If p is the probability of success, q that of failure, then we are given that $p : q = 2 : 1$ providing $p = 2/3$, $q = 1/3$. Hence, for a Bernoulli trial

$$P(X = x) = \binom{n}{x} q^{n-x} p^x = \binom{6}{x}\left(\frac{1}{3}\right)^6 (2)^x.$$

Let $L_4 =$ {At least four successes} so $P(L_4) = P(X \ge 4) = P(X = 4) + P(X = 5) = P(X = 6)$.

$$\therefore \quad (L_4) = \left(\frac{1}{3}\right)^6 \left\{16\binom{6}{4} + 32\binom{6}{5} + 64\binom{6}{6}\right\} = \frac{240 + 192 + 64}{729} = \frac{496}{729}.$$

7*. The process can be split into two Bernoulli's trials ;

First trial. P(head) $= 1/2$, $n = 100$, $x = 50$

$$\therefore \quad P(X = 50) = \binom{100}{50}\left(\frac{1}{2}\right)^{50}\left(\frac{1}{2}\right)^{50} = \frac{100!}{50!\,50!}\left(\frac{1}{2}\right)^{100} \doteq 0.0796.$$

Second Trial. Event of throwing 50 heads. Here $p = 0.0796$, $n = 5$, $r = 2$.

$$\therefore \quad P(R = 2) = \binom{5}{2}(0.9204)^3 (0.0796)^2 = 10 \times 0.78 \times 0.00634) = 0.049.$$

The required probability P is obtained by the principle of sequential counting

$$P = 0.049 \times 0.0796 = 0.039.$$

8*. We consider the sample space S where $S = \{H > T,\ H = T,\ H < T\}$, where $H > T$ means more heads occur than tails and so on. By symmetry, $P(H > T) = P(H < T)$.

Now $P(S) = P\{(H > T) \uplus (H = T) \uplus (H < T)\} \Rightarrow 1 = 2P(H > T) + P(H = T)$.

Thus $$P(H > T) = (1/2)\,[1 - P(H = T)] \qquad \ldots(1)$$

Now $P(H) = P(T) = 1/2$, and to obtain equal number of heads and tails out of $2n$ flips, we require n heads, whence

$$P(X = T) = \binom{2n}{n}\left(\frac{1}{2}\right)^{2n-n}\left(\frac{1}{2}\right)^{n} - \binom{2n}{n}\left(\frac{1}{2}\right)^{2n} \qquad \ldots(2)$$

$$P(H > T) = \frac{1}{2}\left[1 - \binom{2n}{n}\left(\frac{1}{2}\right)^{2n}\right]. \qquad \text{[by (1) and (2)]}$$

9*. Probability of a head turning up on a single toss is $\frac{1}{2}$, and that of k heads in n throws is

$$P(X = k) = {}^nC_k\, p^k q^{n-k} = {}^nC_k \left(\tfrac{1}{2}\right)^n \qquad (p = \tfrac{1}{2})$$

$$\therefore \quad P(X \le r) = \sum_{k=0}^{r} P(X = k) = \left[\binom{n}{0} + \binom{n}{1} + \ldots + \binom{n}{r}\right]\left(\frac{1}{2}\right)^n.$$

10*. Let p denote the probability of success and $q = 1 - p$. Then

$$P(E_r) = {}^nC_r\, p^{n-r} q^r,\ P(E_{n-r}) = {}^nC_{n-r}\, q^{n-r} p^r$$

$$\therefore \quad \frac{P(E_r)}{P(E_{n-r})} = \left(\frac{p}{q}\right)^{n-2-r} \quad \text{which is independent of } n \text{ iff } \frac{p}{q} = 1.$$

Thus $$p = q \Rightarrow p = q = 1/2.$$

11*. Let $p = P\{$even face uppermost on a die$\}$. Then the probability of getting x even numbers in 10 throws is

$$P(x) = \binom{10}{x} p^x q^{10-x},\ x = 0, 1, \ldots, 10. \qquad \ldots(1)$$

We are given that $P(X = 5) = 2P(X = 4) \Rightarrow \binom{10}{5} p^5 q^5 = 2\binom{10}{4} p^4 q^6$

This gives, $(1/5)\, p = (1/3) q \quad \Rightarrow p = 5/8$.

$$\therefore \quad P(x) = \binom{10}{x}\left(\frac{5}{8}\right)^x \left(\frac{3}{8}\right)^{10-x}, \qquad \text{[by (1)]}$$

The required number f when 10,000 sets of 10 throws produce no even number is obviously

$$f = 10{,}000\ P(0) = 10{,}000\ (3/8)^{10} = 1.$$

12*. Let p denote the probability, that a given ball is in one of the first r boxes, so that $p = (r/N)$. Treat the distribution of n balls as an n-fold repetition of the experiment of placing a ball into one of the N boxes. The experiment succeeds with prob. $q = 1 - (r/N)$, when ball does not fall in the first r boxes. The problem is now a Bernoulli's trial; hence the required probability P is given by

$$P = \binom{n}{k} q^{n-k} p^k = \binom{n}{k}\left(1-\tfrac{r}{N}\right)^{n-k}\left(\tfrac{r}{N}\right)^k.$$

13*. (a) A will beat B if A wins three games out of four games played, and the corresponding probability p_1 is

$$p_1 = \binom{4}{3}\left(\tfrac{1}{2}\right)^{4-3}\left(\tfrac{1}{2}\right)^3 = \tfrac{1}{4} \qquad \left(p=q=\tfrac{1}{2}\right)$$

Again, A will beat B if A wins five games out of eight played, and the corresponding probability p_2 is

$$p_2 = \binom{8}{5}\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)^5 = \frac{56}{(2)^8} = \frac{7}{32}.$$

Since $p_1 > p_2$, A is more likely to win 3 games out of 4 than to win 5 games out of 8.

(b) The probability of winning at least 3 games out of 4 is

$$p_3 = \binom{4}{3}\left(\tfrac{1}{2}\right)^4 + \binom{4}{4}\left(\tfrac{1}{2}\right)^4 = \tfrac{1}{4} + \tfrac{1}{16} = \tfrac{5}{16}.$$

The probability of winning at least 5 games out of 8 is

$$p_4 = \left[\binom{8}{5}+\binom{8}{6}+\binom{8}{7}+\binom{8}{8}\right]\left(\frac{1}{2}\right)^8 = \frac{93}{256}.$$

Since $p_4 > p_3$, it is more probable to win at least 5 games out of 8.

14*. Suppose X is the number of persons who make reservations on a flight and do not show up for the flight. Then $X \sim$ bin (100, 0.05). Thus

$$P\{X \le 95\} = 1 - P(X > 95) = 1 - \sum_{x=96}^{100} \binom{100}{k} . (0.05)^k\,(0.95)^{100-k}.$$

15*. The probability that a thing is received by a man is $p = a/(a + b)$. The probability that a thing is received by a women is $q = b/(a + b)$. Hence the probability that $(2r + 1)$ things are received by man is

$$P(X = 2r + 1) = \binom{m}{2r+1} p^{2r+1} q^{m-2r-1},\ r = 0, 1, 2, 3, \ldots$$

The chance that the number of things received by men is odd is

$$P_0 = \Sigma P(X = 2r + 1), r = 0, 1, 2, \ldots$$

$$P_0 = \binom{m}{1} q^{m-1} p + \binom{m}{3} q^{m-3} p^3 + \binom{m}{5} q^{m-5} p^5 + \ldots \qquad \ldots(1)$$

From the two binomial expansions

$$(q+p)^m = q^m + \binom{m}{1} q^{m-1} p + \binom{m}{2} q^{m-2} p^2 + \binom{m}{3} q^{m-3} p^3 + \ldots \qquad \ldots(2)$$

$$(q-p)^m = q^m - \binom{m}{1} q^{m-1} p + \binom{m}{2} q^{m-2} p^2 - \binom{m}{3} q^{m-3} p^3 + \ldots \qquad \ldots(3)$$

Subtracting (3) from (3) and using (1)

$2P_0 = (q+p)^m - (q-p)^m \Rightarrow P_0 = (\frac{1}{2})\{(b+a)^m - (b-a)^m\}/(b+a)^m$, [putting for p and q]

16*. Since n is very large even number, let $n = 2k$, where k itself is very large number. Also p = prob. of a getting a head on a single toss $= \frac{1}{2}, q = \frac{1}{2}$.

The prob. $f(x)$ of getting $(k-x)$ heads and $(k+x)$ tails is given by

$$f(x) = \binom{2k}{k-x} p^{k-x} q^{k+x} = \binom{2k}{k-x}\left(\frac{1}{2}\right)^{2k}.$$

The prob. of getting k heads and k tails (put $x = 0$) is thus $f(0)$.

$$\therefore \frac{f(x)}{f(0)} = \frac{\binom{2k}{k-x}}{\binom{2k}{k}} = \frac{\left(\frac{2k}{k-x}\right)^{k-x+1/2}\left(\frac{2k}{k+x}\right)^{k+x+1/2}}{(2)^{k+1/2}(2)^{k+1/2}} = \left(1+\frac{x}{k}\right)^{k+x+1/2}\left(1-\frac{x}{k}\right)^{k-x+1/2} \quad \text{[by Stirling Formula]}$$

Taking logarithms we obtain

$$\ln\frac{f(x)}{f(0)} = \left(k+x+\frac{1}{2}\right)\ln\left(1+\frac{x}{k}\right) + \left(k-x+\frac{1}{2}\right)\ln\left(1-\frac{x}{k}\right)$$

$$= \left(k+x-\frac{1}{2}\right)\left(\frac{x}{k}-\frac{x^2}{2k^2}+\ldots\right) - \left(k-x+\frac{1}{2}\right)\left(\frac{x}{k}+\frac{x^2}{2k^2}+\ldots\right) = -\left(\frac{x^2}{k}\right) + 0\left(\frac{1}{k^2}\right)$$

This gives : $\quad f(x) = f(0).^{-x^2/k}, \quad$ (since $k \to \infty$).

As $\quad f(0) = \binom{2k}{k}\left(\frac{1}{2}\right)^{2k} = \frac{1}{\sqrt{4\pi k}}(2)^{k+1/2}(2)^{k+1/2}\left(\frac{1}{2}\right)^{2k} = \frac{1}{\sqrt{\pi k}}$

the above result reduces to $f(x) = (2/\pi n)^{1/2}\, e^{-(2x^2/n)}$, as $k = n/2$.

Sec. 12-21. Page 390

1*. Since the number of black balls equals the number of white balls, $p = P$ {a black ball drawn} = 1/2. Thus, $p = q = 1/2$. and $n = 5$. The frequency of x ball's being black is

$$f(x) = 819\binom{5}{x}\left(\frac{1}{2}\right)^{5-x}\left(\frac{1}{2}\right)^x = \frac{819}{32}\binom{5}{x}. \qquad [\because \Sigma f = 819]$$

This gives $f(0) = 819/32 = 25.6$. We now use Iteration formula :

$$f(x+1) = \frac{n-x}{x+1}\frac{p}{q}f(x), \text{ i.e. } f(x+1) = \frac{5-x}{x+1}f(x).$$

$\therefore f(1) = 5f(0) = 128, f(2) = 2f(1) = 256, f(3) = f(2) = 256, f(4) = \frac{1}{2}, f(3) = 128, f(5) = \frac{1}{5}, f(4) = 25.6.$

The following table compares the frequencies :

Observed frequencies :	30	125	277	224	136	27
Expected frequencies :	25.6	128	256	256	128	25.6.

2*. $p = P\{\text{rain falls}\} = 10/30 = 1/3,\ q = 2/3.$

(a) If p_x denotes the prob. that rain falls on x days of a week, then

$$p_x = \binom{n}{x} q^{n-x} p^x = {}^7C_x\, 2^{7-x} / (3)^7, \quad x = 0, 1, 2, 3, 4, 5, 6, 7.$$

$$p_3 = {}^7C_3\, 2^4/3^7 = 35 \times 16 / 2187 = 560 / 2167 = 0.2561.$$

Using $f(x+1) = \dfrac{n-x}{x+!}\dfrac{p}{q} f(x), [n = 7, p/q = \frac{1}{2}]$ we get $f(x+1) = \{(7-x)/2(x+1)\} f(x)$

$p_4 = \frac{1}{2} p_3 = 0.1280,\ p_5 = (3/10),\ p_4 = 0.0384,\ p_6 = (1/6)\, p_5 = 0.0064,\ p_7 = (1/14)\, p_6 = 0.0005.$

$P(L_3) = p_3 + p_4 + p_5 + p_6 + p_7 = 0.4294.$

We now attend to the 2nd part. Let W_i mean ith day is wet, etc. Then

$$P\{W_1 W_2 W_3 W_4 D_5 D_6 D_7\} = P(W_1) P(W_2) P(W_3) P(W_4) P(D_5) P(D_6) P(D_7) = \{P(W)\}^4 \{P(D)\}^3$$

$$= (1/3)^4 (2/3)^3 = 8/2187 = 0.0037.$$

3*. If p is the probability of safe arrival, then we are given $p = 9/10$, $q = 1/10$, $n = 150$.

$$\therefore \quad \text{Mean} = np = 135, \sigma = \sqrt{npq} = \sqrt{13.5} = 3.674.$$

Sec. 12-31. Page 392

1*. (i) $(n + 1)p = 4$, hence $k = 4$ and $k = 3$ are two modal values.

(ii) $[(n + 1)p] = [17/2] = 8$, hence $k = 8$ is the modal value.

2*. Let X be the total gain of the player A. The r.v. X assumes the values $-1, 0, 1, 7$ with the following chances :

$$P(X=-1) = P\{\text{no sixes}\} = \left(\frac{5}{6}\right)^3 = \frac{125}{216};\ P(X=0) = P\{1\text{ six}\} = \binom{3}{1}\left(\frac{1}{6}\right)\left(\frac{5}{6}\right)^2 = \frac{75}{216}$$

$$P(X=1) = P\{2\text{ sixes}\} = \binom{3}{2}\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right) = \frac{15}{216};\ P(X=2) = P\{3\text{ sixes}\} = \binom{3}{3}\left(\frac{5}{6}\right)^0\left(\frac{1}{6}\right)^3 = \frac{1}{216}.$$

$$\therefore \quad \mathsf{E}(X) = (-1)\frac{125}{216} + 0.\frac{75}{216} + 1.\frac{15}{216} + 7\frac{1}{216} = \frac{-103}{216} < 0.$$

Since $\mathsf{E}(X) < 0$, the game is not fair.

The game may become fair if A gets Rs c when three sixes appear and c satisfies the condition

$$0 = \mathsf{E}(X) = \frac{-125}{216} + 0 + \frac{15}{216} + \frac{c-1}{216} \Rightarrow c = \text{Rs } 111.$$

3*. Here $\mu = np$. Chebyshev's Inequality : $P\{\,|X - \mu| \ge a\} \le \sigma^2/a^2$.

(a)
$$p = P\{|X - 10| < 8\} \ge 1 - \sigma^2/64 = 1 - 15/128 = 113/128.$$

$$p = P\{|X - 10| > 10\} < \frac{15/2}{100} = 0.075.$$

(b) $p = P\{|X-20|\leq 5\} = 1 - P\{|X-20|>5\} \geq 1 - 10/25 = 3/5.$

(c) $p = P\{|X-20|>10k\} \leq 10/100\,k^2 = \frac{1}{4} \Rightarrow k = 2/\sqrt{10}.$

Sec. 12-43. Page 397

1*. Here $np = 5$, $npq = 9$. If we divide out, we get $q = 9/5 > 1$. This being impossible, the statement is *wrong*.

2*. If σ^2 is the variance, then $\sigma^2 = npq = n(p-p^2)$

$$\frac{d\sigma^2}{dp} = n(1-2p), \frac{d^2(\sigma^2)}{dp^2} = -2n < 0. \text{ So } \frac{d\sigma^2}{dp} = 0 \Rightarrow p = 1/2; \text{ hence } q = 1/2.$$

Thus max (σ^2) = max $(npq) = n/4$.

3*.
$$P(X\geq 1) = 1 - P(X<1) = 1 - P(X=0) = 1 - {}^nC_0\,q^n = 1 - q^n$$
$$= 1-(1-p)^n = 1-(1-np+\ldots) = np = \mu.$$

4*. $y = f(x) = \binom{n}{x} q^{n-x} p^x, \quad [x \neq 0,\ n]$

$$\frac{dy}{dp} = \binom{n}{x}[xp^{x-1}q^{n-x} - (n-x)q^{n-x-1}p^x] = \binom{n}{x} p^x q^{n-x}\left[\frac{x}{p} - \frac{n-x}{q}\right] = y\left[\frac{x}{p} - \frac{n-x}{q}\right].$$

$$\frac{d^2y}{dp^2} = y\left\{\left(\frac{x}{p} - \frac{n-x}{q}\right)^2 + \left[-\frac{x}{p^2} - \frac{n-x}{q^2}\right]\right\}$$

$$\frac{dy}{dp} = 0 \Rightarrow x = np \text{ or } p = \frac{x}{n};\ \left(\frac{d^2y}{dp^2}\right)_{np=x} = -y\left[\frac{x}{p^2} + \frac{n-x}{q^2}\right] = -\frac{ny}{pq} < 0.$$

Hence the estimate (x/n) is the maximum likelihood estimate (brief : MLE) of p because of the above maximizing property.

5*.
$$E\left[\sin\left(\frac{\pi X}{2}\right)\right] = \sum_{x=0}^{4} \sin\left(\frac{\pi x}{2}\right)\binom{4}{x} p^x q^{4-x} = \binom{4}{1} pq^3 - \binom{4}{3} qp^3$$
$$= 4pq(q^2-p^2) = 4pq\,(q-p)\,(q+p) = 4p\,(1-p)\,(1-2p).$$

6*. Here $np = 6$, $npq = 2$. Dividing the latter by former we get $q = 1/3 \Rightarrow p = 2/3$ and $n = 9$. Thus

$$P(X=r) = \binom{9}{r}\left(\frac{1}{3}\right)^{9-r}\left(\frac{2}{3}\right)^r\left(\frac{1}{3}\right)^9 = \binom{9}{r} 2^r, (0 \leq r \leq 9).$$

7*. We are given that : $np = 4$, $npq\ (q-p) = 1.92$.

$\therefore \quad 4q(2q-1) = 1.92 \Rightarrow 2q^2 - q - 0.48 = (q-0.8)\,(2q+0.6) = 0 \Rightarrow q = 0.8.$

Notice that $q \neq -0.3$ as $q \not< 0$. Incidently, $p = 0.2$ and $n = 20$.

$$\text{Var}\,(X) = npq = 3.2 \text{ giving } \sigma = \sqrt{3.2} = 1.789.$$
$$\mu_4 = npq\,\{1 + 3pq\,(n-2)\} = 3.2\,[1 + 9.6 - 0.96] = (3.2)\,(9.64) = 30.848.$$

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{(1.92)^2}{(3.2)^3} = 0.1125;\ \gamma_1 = \sqrt{\beta_1} = 0.3354,\ \beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{20.848}{(3.2)^2} = 3.0125;\ \gamma_2 = \beta_2 - 3 = 0.0125.$$

Sec. 12-63. Page 402

1*. By additive property : $(X_1 + X_2)$ is bin $(n_1 + n_2, p)$ whenever X_i is bin (n_i, p). Now

$$P(\overline{X} = 3/5) = P(X_1 + ... + X_5 = 3) \sim \text{bin}\ (5, 1/2)$$

$$\therefore \quad P(Z = 3) = \binom{5}{3}\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)^2 = \frac{10}{32}, \quad [Z = X_1 + ... + X_5 \sim \text{bin}\ (5, 1/2)].$$

$$\text{(b)}\ P(X > 3/5) = P(Z>3) = P\{(Z=4) \uplus (Z=5)\} = p_4 + p_5 = \binom{5}{4}\left(\frac{1}{2}\right)^5 + \binom{5}{5}\left(\frac{1}{2}\right)^5 = \frac{6}{32}.$$

$$\text{(c)}\ P(\overline{X} < 3/5) = P(Z < 3) = P\{(Z = 0) \uplus (Z = 1) \uplus (Z = 2)\} = p_0 + p_1 + p_2$$

$$= \left[\binom{5}{0} + \binom{5}{1} + \binom{5}{2}\right]\left[\frac{1}{2}\right]^5 = \frac{(1 + 5 + 10)}{32} = \frac{1}{2}.$$

2*. For fixed value of $P = p$, say. $G_X(t) = (q + pt)^n$

Differention gives $\quad G_X^r(t) = n^{(r)}p^r(q + pt)^{n-r} \Rightarrow \mathrm{E}(X^{(r)}) = G^r(1) = n^{(r)}p^r$.

Now, we use Double-E Rule.

$$\mathrm{E}(X^{(r)}) = \mathrm{E}_P[\mathrm{E}_X(X^{(r)} \mid P = p)] = \mathrm{E}_P[n^{(r)}P^r] = n^{(r)}\,\mathrm{E}(P^r) = n^{(r)}\,\mu_r'.$$

3*. $\qquad G(t : X_i) = t^1 p + t^0 q = q + pt$

$G(t : X_1 + X_2 + ... + X_n) = G(t : X_1)\, G(t : X_2) \ldots G(t : X_n) \qquad [\because \text{Ind}\ (X_i, X_j)]$

$$\therefore \quad G(t : X) = (q + pt)^n = \sum_{x=0}^{n}\binom{n}{x}q^{n-x}(pt)^x,\ x = 0, 1, ..., n.$$

Thus $\qquad P(X = x) = \text{Coeff. of } t^x \text{ in } G(t : X) = \binom{n}{x}q^{n-x}p^x,\ (0 \le x \le n).$

Remark. This Example provides another derivation of Binomial Distribution.

4*. Here $\qquad G(t : X_i) = \Sigma p_i t^{X_i} = pqt^0 + (p^2 + q^2)t^1 + pqt^2 = pq + (p^2 + q^2)t + pqt^2.$

$$G(t : \Sigma X_i) = [G(t : X_i)]^n = [pq + (p^2 + q^2)t = pqt^2]^n \qquad [X_k \text{ are i.i.d.}] \qquad ...(1)$$

Since $X \sim$ bin (n, p), $Y \sim$ bin (n, q) and Ind. (X, Y) we have

$$G(t : X) = (1 - p + pt)^n, \quad G(t : Y) = (1 - q + qt)^n$$

$\therefore \quad G(t : X + Y) = G(t : X).G(t : Y) = (q + pt)^n (p + qt)^n = [pq + (p^2 + q^2)t + pqt^2]^n$ ...(2)

From (1) and (2), $G(t : \Sigma X_i) = G(t : X + Y)$, hence the result.

5*. $P\{X = x \mid X > 0\} = P\{X = x\}/P(X > 0) = f(x)/[1 - P(X = 0)] = f(x)/(1 - q^n)$

$$G(t : X) = \mathrm{E}(t^X) = \sum_{x=1}^{n} t^x \left[\frac{f(x)}{1-q^n}\right] = \sum_{x=1}^{n} t^x \cdot \frac{\binom{n}{x} q^{n-x} p^x}{(1-q^n)}$$

$$\therefore \quad (1 - q^n)\, G_X(t) = \sum_{x=0}^{n} t^x f(x) - f(0) = (q + pt)^n - q^n \qquad [\text{Added} : f(0) - f(0)]$$

Thus $\quad G(t) = [(q + pt)^n - q^n]/(1 - q^n).$ ...(1)

Differentiating (1), k times w.r.to t yields

$(1 - q^n)G^{(k)}(t) = n^{(k)} p^k (q + pt)^{n-k} \quad \Rightarrow \quad \mathrm{E}[X^{(k)}] = G^{(k)}(1) = n^{(k)} p^k / (1 - q^n)$ ...(2)

So $\quad \mathrm{E}(X) = np/(1 - q^n)$; $\mathrm{E}[X(X-1)] = n(n-1)\, p^2 (1 - q^n)$.

$$\mathrm{Var}(X) = \mathrm{E}(X^2) - \mathrm{E}^2(X) = \frac{n(n-1)\, p^2}{(1-q^n)} + \frac{np}{1-q^n} - \left(\frac{np}{1-q^n}\right)^2 = \frac{npq}{1-q^n}\left(1 - \frac{npq^{n-1}}{1-q^n}\right).$$

Sec. 12-80. Page 409

1*. For the binomial frequency distribution, $N\left(\frac{1}{2} + \frac{1}{2}\right)^n$, we have

$$N\left(\frac{1}{2} + \frac{1}{2}\right)^n = N\left\{\binom{n}{0}\left(\frac{1}{2}\right)^n + \binom{n}{0}\left(\frac{1}{2}\right)^n + \ldots + \binom{n}{n}\left(\frac{1}{2}\right)^n\right\}$$

or
$$N = N\left(\frac{1}{2}\right)^n \left\{\binom{n}{0} + \binom{n}{1} + \ldots + \binom{n}{r-1} + \binom{n}{r} + \ldots + \binom{n}{n}\right\} \qquad \ldots(1)$$

The other symmetrical distribution is identical with (1). Hence we superimpose the two distributions, as *suggested*, to obtain

$$2N = N\left(\frac{1}{2}\right)^n \left\{\binom{n}{0} + \left[\binom{n}{0} + \binom{n}{1}\right] + \ldots + \left[\binom{n}{r-1} + \binom{n}{r}\right] + \ldots + \left[\binom{n}{n-1} + \binom{n}{n}\right] + \binom{n}{n}\right\}$$

or
$$N = N\left(\frac{1}{2}\right)^{n+1} \left\{\binom{n+1}{0} + \binom{n+1}{1} + \ldots + \binom{n+1}{r} + \ldots + \binom{n+1}{n} + \binom{n+1}{n+1}\right\} \qquad \ldots(2)$$

because
$$\binom{n}{r-1} + \binom{n}{r} = \binom{n+1}{r},\ \binom{n}{0} = \binom{n+1}{0};\ \binom{n}{n} = \binom{n+1}{n+1}.$$

Comparing (1) and (2) we see that the superimposed distribution is also symmetrical, with degree $(n + 1)$.

2*. Let $A = \{X = n \mid X > 0\}$, $B_j = \{p = p_j\}$.

Then by Multi-Stage p-Rule : $P(A) = P(B_1)\, P(A \mid B_1) + P(B_2)\, P(A \mid B_2)$ ...(1)

Now $P(A \mid B_1) = P\{X = n, p = p_1 \mid X > 0\} = \dfrac{P(X = n)}{P(X > 0)} = \dfrac{\binom{n}{n} q_1^{n-n} p_1^n}{(1 - q_1^n)} = \dfrac{p_1^n}{1 - q_1^n}$.

Similarly, $P(A \mid B_2) = p_2^n / (1 - q_2^n)$. Since $P(B_1) = \theta$, $P(B_2) = 1 - \theta$, we readily obtain from (1)

$$P(A) = \frac{\theta\, p_1^n}{1 - q_1^n} + \frac{(1-\theta)\, p_2^n}{1 - q_2^n}.$$

3*. Here, probability of success $p = 0.25$, $q = 0.75$. The sample size, or number of runs n is not known.

$$(q + p)^n = q^n + \binom{n}{1} q^{n-1} p + \binom{n}{2} q^{n-2} p^2 + \ldots$$

$(0.75 + 0.25)^n = (0.75)^n + n(0.75)^{n-1} \times (0.25) + \ldots +$ [$(q + p)^n$ gives various probs.] ...(1)

These first two terms represent the prob. of 0 success and 1 success respectively. the remaining terms give respectively probabilities of 2, 3, 4, ..., successes. We want that these latter probabilities account for 80% of the possible outcomes, hence the sum of the first two terms in (1) must be no more than 0.20. Thus

$(0.75)^n + n(0.75)^{n-1} \times (0.25) \le 0.20$, i.e. $(0.75)^{n-1} [(0.75) + n \times (0.25)] \le 0.20$

or $\quad (0.75)^{n-1} (3 + n) \le 0.80$.

We solve this equation for n, by trial and error. Thus :

$n = 5$, L.H.S. $= 2.5312$; $n = 10$. L.H.S. $= 0.9764$; $n = 11$, L.H.S. $= 0.7884$.

This shows that a minimum of 11 runs will give at least an 80% assurance of at least two successes.

4*. Let X and Y denote the number of successes in the first half and in the second half of n events respectively. Then $X \sim \text{bin}\left(\frac{1}{2}n, p\right)$ $Y \sim \text{bin}\left(\frac{1}{2}n, q\right)$.

$$\therefore \qquad E(X) = np/2,\ \text{Var}(X) = npq/2;\quad E(Y) = nq/2,\ \text{Var}(Y) = npq/2.$$

If Z denotes all the successes in n events, then $Z = X + Y$, so that

$$E(Z) = E(X + Y) = E(X) + E(Y) = n(p + q)/2 = n/2.$$

$$\text{Var}(Z) = \text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) = (npq/2) + (npq/2) = npq.$$

Notice that X and Y are independent variates.

5*. We shall utilize mathematical induction to establish the given result. Note also that $q =$ const. during the course of differentiation w.r.t. p. Now

$$T \equiv (q + p)^n = \sum_{x=0}^{n} \binom{n}{x} q^{n-x} p^x;\quad \frac{\partial T}{\partial p} = \sum_{x=0}^{n} \binom{n}{x} q^{n-x} p^{x-1} . x.$$

$$\therefore \qquad p\frac{\partial T}{\partial p} = p \sum_{x=0}^{n} \binom{n}{x} q^{n-x} p^{x-1} x = \sum_{x=0}^{n} \binom{n}{x} q^{n-x} p^x x = \mu_1'.$$

It follows that the given result is true for $k = 1$. This starts Induction. We assume that the given result is true for $1 \le k = m$; so that

$$\left(p\frac{\partial}{\partial p}\right)^m T = \mu'_m = \sum_{x=0}^{n}\binom{n}{x} q^{n-x} p^x x^m.$$

Differentiate this w.r.t. 'p' and multiply both sides with p, to obtain

$$\left(p\frac{\partial}{\partial p}\right)\left[\left(p\frac{\partial}{\partial p}\right)^m T\right] = \sum_{x=0}^{n}\binom{n}{x} q^{n-x} p^x x^{m+1} \Rightarrow \left(p\frac{\partial}{\partial p}\right)^{m+1} T = \mu'_{m+1}.$$

It follows by mathematical induction that if the result is true for $k = m$, it is also true for $k = m + 1$. It is already shown to be true for $k = 1$, hence the result is true for any integer k.

Note. For $k = 2$, we have

$$\mu'_2 = p\frac{\partial}{\partial p}\{np(q+p)^{n-1}\} = np\{(q+p)^{n-1} + (n-1)\,p(q+p)^{n-2}\} = np(1-p+np) = npq + (np)^2$$

and so on. Central moments and β_1, β_2 can thus be calculated.

6*. By Partition Theorem : $P(X = x) = P(X = x, Y = 0) + P(X = x, Y = 1)$ hence

$$P(X = x) = \left(\frac{(1-\theta)\theta'}{\theta+\theta'}\right)^{1-x}\left(\frac{\theta\theta'}{\theta+\theta'}\right)^{x} + \left(\frac{\theta\theta'}{\theta+\theta'}\right)^{1-x}\left(\frac{\theta(1-\theta')}{\theta+\theta'}\right)^{x}$$

$$= [\theta^x(1-\theta)^{1-x}\,\theta' + \theta(\theta')^{1-x}(1-\theta')^x]/(\theta+\theta')$$

Thus, $$P(X = 1) = \frac{\theta\theta' + \theta(1-\theta')}{\theta+\theta'} = \frac{\theta}{\theta+\theta'} = p\,(\text{say})\ ;\quad P(X=0) = \theta'/(\theta+\theta') = q.$$

Thus, $X \sim \text{bin}\,(1, p)$, where $p = \theta/(\theta + \theta')$. By symmetry, $Y \sim \text{bin}\,(1, p)$. $[f(x, y) = f(y, x)]$

$$\therefore \quad \mathrm{E}(X) = \mathrm{E}(Y) = p,\ \mathrm{Var}\,(X) = \mathrm{Var}\,(Y) = pq.$$

Further, from given p.m.f. $\mathrm{E}(XY) = P(X = 1, Y = 1) = \theta(1 - \theta')/(\theta + \theta')$.

$$\therefore \quad \sigma_{XY} = \mathrm{E}(XY) - \mathrm{E}(X)\,\mathrm{E}(Y) = \frac{\theta(1-\theta')}{\theta+\theta'} - \left(\frac{\theta}{\theta+\theta'}\right)^2 = \frac{\theta\theta'}{(\theta+\theta')^2}(1-\theta-\theta')$$

$$\mathrm{Var}\,(X) = pq = \theta\theta'/(\theta+\theta')^2 . \rho(XY) = (\sigma_{XY}/\sigma_X . \sigma_Y) = (1-\theta-\theta').$$

Chapter 13 : *Poisson Distribution*

Sec. 13-12. Page 417

1*. Here $P(X = r) \equiv f(r) = e^{-\lambda}\lambda^r/r!,\ \ r = 0, 1, 2, \ldots$

(a) $$f(1) = f(2) \Rightarrow \lambda e^{-\lambda} = \tfrac{1}{2}e^{-\lambda}\lambda^2 \Rightarrow \lambda = 2 = \mathrm{E}(X)$$

Thus $$f(4) = e^{-2}2^4/4! = 2e^{-2}/3 = 0.0902.\ \ [e^{-2} = 0.13534]$$

(b) $$2f(0)+f(2) = 2f(1) \Rightarrow 2e^{-\lambda}+(e^{-\lambda}\lambda^2/2)=2\lambda e^{-\lambda}$$

Thus $$\lambda^2-4\lambda+4 = 0 = (\lambda-2)^2 \Rightarrow \lambda=\mathrm{E}(X)=2.$$

(c) $f(2) = 9f(4)=90f(6) \Rightarrow \left(\frac{1}{2}\right)\lambda^2 = 9(\lambda^2/24)+90(\lambda^6/30\times 24),$ [cancel $e^{-\lambda}$]

$\therefore$ $\lambda^4+3\lambda^2-4=0=(\lambda^2-1)(\lambda^2+4) \Rightarrow \lambda=1.(\lambda \not< 0).$

(d) Here $f(k) = e^{-1}/k!,\ k=0,1,2,\dots$ Now the required probability p is

$$p = P(X\geq 2 \mid X\leq 4)=\frac{P(2\leq X\leq 4)}{P(X\leq 4)}=\frac{f(2)+f(3)+f(4)}{f(0)+f(1)+\dots+f(4)}=\frac{e^{-1}\left[\left(\frac{1}{2}\right)+\left(\frac{1}{6}\right)+\left(\frac{1}{24}\right)\right]}{e^{-1}\left[1+1+\left(\frac{1}{2}\right)+\left(\frac{1}{2}\right)+\left(\frac{1}{24}\right)\right]}$$

2*. $$P(X\geq 2) = 1-[P(X=0)+P(X=1)]=1-e^{-\lambda}(1+\lambda).$$

Also $\int_0^{\lambda} xe^{-x}\,dx\left[x(e^{-x})-e^{-x}\right]_0^{\lambda}=1-e^{-\lambda}(1+\lambda).$

Equating the two results, we obtain (1).

3*. $$P\{X=x \mid X\geq 2\} = \frac{P(X=x)}{P(X\geq 2)}=\frac{f(x)}{1-P(X\leq 1)},\quad x=2,3,4,\dots$$

$$g(x) = \frac{e^{-m}\,m^x/x!}{1-(e^{-m}+me^{-m})}=\frac{m^x}{x![e^m-1-m]},\quad x=2,3,4,\dots$$

4*. Recall decreasing factorial power : $n^{(k)} = n!/(n-k)!$. Now

$$\sum_{n=r}^{\infty} n^{(k)}\,P(X=n) = \sum_{n=r}^{\infty}\frac{n!}{(n-k)!}\,\frac{e^{-\lambda}\lambda^n}{n!}=(\lambda)^k\sum_{n=r}^{\infty}\frac{e^{-\lambda}\lambda^{n-k}}{(n-k)!}\qquad [\text{Pit } n-k=t]$$

$$= \lambda^k\sum_{t=r-k}^{\infty}\frac{e^{-\lambda}\lambda^t}{t!}=\lambda^k\,P\{X\geq(r-k)\}.$$

5*. Here $$\frac{P(r+j)}{P(r+j-1)} = \frac{e^{-\lambda}\lambda^{r+j}}{(r+j)!}\,\frac{(r+j-1)!}{e^{-\lambda}\lambda^{r+j-1}}=\frac{\lambda}{r+j}\leq\frac{\lambda}{r+1}.$$

$$P(X=r+k) = P(r).\frac{P(r+1)}{P(r)}.\frac{P(r+2)}{P(r+1)}\cdots\frac{P(r+k)}{P(r+k-1)}\leq P(r)\left(\frac{\lambda}{r+1}\right)^k,\quad 0\leq k<\infty.$$

If $\lambda<r+1$, then summing over k, we get

$$\Sigma_k\,P(X=r+k) \leq \frac{P(r)}{1-[\lambda/(r+1)]} \Rightarrow \frac{P(r)}{\Sigma_k\,P(r+k)}\geq 1-\frac{\lambda}{r+1}.$$

Now, as $r\to\infty$. $\lim\limits_{r\to\infty}\dfrac{P(r)}{\Sigma_k\,P(r+k)}=\lim\limits_{r\to\infty}P\{X=r \mid X\geq r+k\}=1.$

6*. Here $$\frac{P(r+j)}{P(r+j-1)} = \frac{e^{-\lambda}\lambda^{r+j}}{(r+j)}\,\frac{(r+j-1)!}{e^{-\lambda}\lambda^{r+j-1}}=\frac{\lambda}{r+j}\leq\frac{\lambda}{r+1}.$$

$$P(X=r+k) = P(r).\frac{P(r+1)}{P(r)}.\frac{P(r+2)}{P(r+1)}\cdots\frac{P(r+k)}{P(r+k-1)}\leq P(r)\left(\frac{\lambda}{r+1}\right)^k,\quad 0\leq k<\infty.$$

If $\lambda < r+1$, then summing over k, we get

$$\Sigma_k P(X = r+k) \le \frac{P(r)}{1-[\lambda/(r+1)]} \Rightarrow \frac{P(r)}{\Sigma_k P(r+k)} \ge 1 - \frac{\lambda}{r+1}.$$

Now, as $r \to \infty$, $\lim\limits_{r\to\infty} \dfrac{P(r)}{\Sigma_k P(r+k)} = \lim\limits_{r\to\infty} P\{X = r \mid X \ge r+k\} = 1$

Remarks. If r is large enough, and if we have enjoyed r trouble-free hours, then it is almost certain that there will be a breakdown during the next hour.

7*. Here $S = P(X \ge k) + P(X \ge k+1) + P(X \ge k+2) \ldots = \left\{ \sum\limits_{x=k}^{\infty} f(x) + \sum\limits_{x=k+1}^{\infty} f(x) + \sum\limits_{x=k+2}^{\infty} f(x) + \ldots \right\}$

$$= f(k) + 2f(k+1) + 3f(k+2) + \ldots = \Sigma[(k+r)-(k-1)]\, f(k+r), 0 \le r < \infty.$$

$$= (1-k)\sum_{r=0}^{\infty} \frac{e^{-\lambda}\lambda^{k+r}}{(k+r)!} + \lambda \sum_{r=0}^{\infty} \frac{e^{-\lambda}\lambda^{k+r-1}}{(k+r-1)!} = (1-k)\sum_{t=k}^{\infty} \frac{e^{-\lambda}(\lambda)^t}{t!} + \lambda \sum_{t=k-1}^{\infty} \frac{e^{-\lambda}(\lambda)^t}{t!}$$

$$= (1-k)\, P(X \ge k) + \lambda P\,(X \ge k-1).$$

8*. Y = bin (x, p) and $X \sim$ pois (λ); hence using Double-E Rule we have

$$G(t:Y) = \mathrm{E}(t^Y) = \mathrm{E}\{\mathrm{E}(t^Y \mid X = x)\} = \mathrm{E}\{(q+pt)^X\} = \mathrm{E}(t_1)^X.\,[\text{p.g.f. of Pois }(\lambda)]$$

$$= e^{\lambda(t_1 - 1 = e^{\lambda}(q+pt-1)} = e^{\lambda p(t-1)}.$$

Thus $Y \sim$ Pois (λp).

Sec. 13-21. Page 421

1*. Let the variate X denote the number of individuals suffering from a bad reaction. Assuming bad reactions to be rare, we treat X as Poisson rather than Bernoulli's variate. Now m = np = (2000) (0.001) = 2 ; hence

$$P(X=3) = e^{-2}\,2^3/3! = 4 \times 0.1353352/3 = 0.180447.$$

$$P(X>2) = 1 - P(X \le 2) = 1 - [P(X=0) + P(X=1) + P(X=2)]$$

$$= 1 - e^{-2}[1+2+2] = 1 - 5 \times 0.1353352 = 1 - 0.6766764 = 0.3233235.$$

2*. Here p = 0.01, n = 100, np = 1 = m. If X is the number of defective items in a sample of 100, then we need find $P(X = 1)$. Here

$$P(X=0) = e^{-1},\ P(X=1) = e^{-1};\ \text{hence } P(X \le 1) = P(X=0) + P(X=1) = 2e^{-1} = 2/e.$$

3*. Here n = 2000, p = 0.0005, hence $m = np$ = 1. Let the r.v. X denote the number of failures. So $P(X \ge 1) = 1 - f(0) = 1 - e^{-1} = 1 - 0.36788 = 0.63$.

4*. King's sample consists of n coins and the probability p of any coin being false is k/n. The drawings being independent, Bernoulli's distribution is applicable. Hence, if X denotes the No. of the false coins, then

$$P(X=r) = \binom{n}{r} q^{n-r} p^r = \binom{n}{r}\left(1-\frac{k}{n}\right)^{n-r}\left(\frac{k}{n}\right)^r.$$

If n is very large, Poisson approximation holds : $P(X=r)=e^{-k}k^r/r!$.

(ii) P\{Ministers' peculations go undetected\} = P\{0 false coin\} = $\binom{n}{0}\left(1-\frac{k}{n}\right)^n=\left(1-\frac{k}{n}\right)^n$.

If n is very large, Poisson approximation : $P(X = r) = \lim\,(1 - k/n)^n = e^{-k}$.

5*. On an average, there is one misprint per page. Hence $\lambda = 1$. Let X represent the misprints per page ; then

$$P(X\geq 4) = 1-P(X\leq 3)=1-\sum_{r=0}^{3}\frac{e^{-\lambda}\lambda^r}{r!}=1-\frac{1}{e}\left[1+1+\frac{1}{2}+\frac{1}{6}\right]$$
$$= 1-(0.36788)(8/3)=1-0.981016=0.02.$$

6*. Out of 100, the number of dead is 0.005, hence mean number of dead per 10,000 is $\lambda = 0.5$. If r.v. X denotes the number of claims, then

$$p = P(X>3)=1-P(X\leq 3)=1-\Sigma\, e^{-m}\, m^x/x!,\quad (0\leq x\leq 3).$$

7*. For Poisson Distribution : $P(X > 0) = 1 - P(X = 0) = 1 - e^{-\lambda}$. Since $\lambda = np$, this gives on taking logarithm : $np = -\ln[1-P(X>0)]$..

In the problem quoted : $p = 0.0002$, $P(X > 0) = 0.95$

$$\therefore \quad n = \frac{-\ln(1-0.95)}{(0.0002)}=(\ln_e 20)\times 5000=(2.3026)(\ln_{10} 20)\,5000$$
$$= (2.3026)(1.3010)(5000)=14978.413=14979 \text{ cubic feet.}$$

8*. Let A = \{no criminal is hanged\}. Then by Multi-Stage Rule,

$$P(A)=\sum_{r=1}^{4}P\{A\mid\lambda=r\}\,P(\lambda=r)=(e^{-1}/3)+(e^{-2}/3)+(e^{-3}/4)+(e^{-4}/12)$$

9*. Let X be the No. of cars crossing the intersection A between 6 P.M. and 7 P.M. by hypothesis, $X \sim$ Pois (2). Now

$$p = P\{\text{Ist car crosses } A \text{ within 2 min.}\}=1-P\{\text{no car crosses } A \text{ within 2min.}\}$$
$$P(T<1)=1-P(X=0)=1-e^{-2}.$$

Remark. We could use Exponential distribution as well.

Sec. 13-40. Page 425

1*. We twice differentiate w.r.t. "λ" the identity $e^{\lambda} = \Sigma(\lambda^r/r!)$, $0\leq r<\infty$ wo obtain.

$$\left\{e^{\lambda}=\sum_{r=0}^{\infty}\frac{r\lambda^{r-1}}{r!},\ e^{\lambda}=\sum_{r=0}^{\infty}\frac{r(r-1)\lambda^{r-2}}{r!}\right\}\Rightarrow\left\{\lambda=\sum_{r=0}^{\infty}\left(\frac{e^{-\lambda}\lambda^x}{r!}\right)r;\ \lambda^2=\sum_{r=0}^{\infty}\left(\frac{e^{-\lambda}\lambda^r}{r!}\right)r(r-1)\right\}$$

These give $\lambda = E(X)$, $\lambda^2 = E(X^2 - X)$; so Var $(X) = E(X^2) - E^2(X) = \lambda$.

$$\boxed{\text{Mean} = \text{variance} = \text{parameter } \lambda}$$

This is a remarkable property of a Poisson variate and is useful for discerning Poissonian nature of a vairate, for example see §13-60 (2) and example 13-33(b).

2*. Here $\lambda = 5$, $\sigma = 4$, so that $\sigma^2 = 16$. Now for a Poisson variate $\lambda = \sigma^2$, the given statement $16 = 5$ is obviously wrong.

3*. Since X is Pois (100), $E(X) = Var(X) = 100$. Let $Z = (X - \mu/\sigma$, i.e. $X = 100 + 10Z$

$$\therefore \quad p = P\{75 < X < 125\} = P(75 < 100 + 10Z \leq 125) = P\{-2.5 < Z < 2.5\}$$
$$= P\{|Z| < 2.5\} \geq 1 - (100/25 \times 25) = 21/25.$$

Thus $p \geq 0.84 \quad \Rightarrow$ lower bound for p is 0.84.

Sec. 13-62. Page 432

1*. We know that if X is Pois (λ), then mean $= \lambda = \sigma^2$. And if X is unimodal, the model measure of skewness S_k is

$$S_k = \frac{\mu - M}{\sigma} = \frac{\lambda - [\lambda]}{\sqrt{\lambda}} \geq 0 \quad (\because \lambda > 0)$$

Since $S_k > 0$, Poisson distribution is skewed to the right.

Note. Recall : The asymmetry (skewness) coefficient of a Pois (λ) is the quotient $S_k = \mu_3/\sigma^3 = \lambda/(\lambda)^{3/2} = 1/\sqrt{\lambda} > 0$. Hence the conclusion : right-skewed follows as before.

2*. The first four moments of the binomial distribution are

$$\mu = np,\ \mu_2 = npq,\ \mu_3 = npq\ (q - p),\ \ \mu_4 = npq\ (1 + 3npq - 6pq) \qquad ...(1)$$

The limit of Binomial distribution, called Poisson's, is obtained by letting $p \to 0$, i.e. $q \to 1$, $n \to \infty$, but $np = \lambda$ (finite). Using the same letters for the moments of Poisson's distribution, we obtain

$$\mu = \lambda = np, \qquad \mu_2 = (np)q = \lim (np) \lim q = \lambda 1 = \lambda.$$
$$\mu_3 = \lim (np) \lim q\ (\lim q - \lim p) = \lambda 1\ (1 - 0) = \lambda$$
$$\mu_4 = \lim (np) \lim q\{1 + 3 \lim (np) \lim q - 6)\ (\lim p)\ (\lim q)\} = \lambda(3 + 3\lambda 1 - 6.0) = \lambda + 3\lambda^2.$$

3*. Let $Y = X_1 + X_2 + ... + X_n = n\overline{X}$. By Reproductive property, Y is Pois ($n\lambda$). Hence

$$P\{\overline{X} = k\} = P\{Y = nk\} = e^{-\lambda n}(\lambda n)^{nk}/(nk)!, \quad nk = 0, 1, 2, ...$$

Thus $\qquad P(\overline{X} = k\} = e^{-\lambda n}(\lambda n)^{nk}/(nk)!, \quad k = 0, 1/n, 2/n, ... \infty.$

4*. Since X_1 and X_2 are independent, we have $M(t : X_1 + X_2) = M\ (t : X_1)\ .\ M(t : X_2)$ ...(1)

As X_1 and $X_1 + X_2$ are Poisson variates, we get from (1)

$$e^{(\lambda + \mu)(e^t - 1)} = e^{\lambda(e^t - 1)} . M(t : X_2) \Rightarrow M(t : X_2) = e^{\mu(e^t - 1)}.$$

The correspondence between m.g.f.s and the distribution functions now reveal that X_2 is Poisson (μ).

5*. $$M_X(t) = E(e^{tX}) = E_N\{E(e^{tX} \mid N)\} = E[e^{N(e^t - 1)}] \qquad \text{[Double-E Rule]}$$
$$= E(e^{t'N}) = M_N(t') = M_N(e^t - 1) \equiv M(e^t - 1 : N)$$

6*. Suppose $f_n(x)$ is the density of $X \mid n$. Then the density $g(x)$ of r.v. X is given by

$$g(x) = \sum_{n=0}^{\infty} f_n(x)\, P(N=n) = \sum_{n=0}^{\infty} f_n(x) \frac{e^{-\theta}\theta^n}{n!}. \qquad \text{[Multi-Stage } p\text{-Rule]}$$

$$\lambda(t) = M(t : X_1 + \ldots + X_n) = M_1(t) \,.\, M_2(t) \ldots M_n(t) \qquad \left[M_i(t) = \mathrm{E}(e^{tX_i})\right] \qquad \ldots(1)$$

Obviously : $M(t : X \mid n) = \Sigma e^{tx} f_n(x)$, and $M(t : X) = \Sigma e^{tx}\, g(x)$.

Now $\quad M_X(t) = \mathrm{E}(e^{tX}) = \mathrm{E}_N\{\mathrm{E}(e^{tX} \mid N)\} = \mathrm{E}_N\{\lambda_N(t)\} = \sum_{n=0}^{\infty} f_n(x) \dfrac{e^{-\theta}\theta^n}{n!} \lambda_n(t) \qquad \ldots(2)$

If X_j are i.i.d variates, then $\lambda(t) = [M(t)^n]$, and we get

$$M_X(t) = \sum_{x=0}^{\infty} \frac{e^{-\theta}\theta^n}{n!} [M(t)]^n = e^{-\theta} \sum_{n=0}^{\infty} \frac{[\theta \,.\, M(t)]^n}{n!} = \exp\{\theta[M(t) - 1]\}$$

7*. Let $e^t = T$(temporarily), then given m.g.f. is

$$M(t : X) = e^{-\lambda} e^{\lambda T} = \sum_{x=0}^{\infty} \frac{e^{-\lambda}(\lambda T)^x}{x!} = \sum_{x=0}^{\infty} \frac{e^{-\lambda}\lambda^x}{x!} e^{tx} = \sum_{x=0}^{\infty} f(x)\, e^{tx}. \text{ (say)} \qquad \ldots(1)$$

By Dirichlets' form of m.g.f. it follows that

$$f(x) = \frac{e^{-\lambda}\lambda^x}{x!}, \quad x = 0, 1, 2, \ldots, \quad \text{i.e. } X \sim \text{Pois}(\lambda).$$

8*. Recalling the m.g.f. of a Pois (λ), we at once see that X is Pois (4). Now for a Pois (λ) $\lambda = \sigma^2 = 4$, hence

$$p = P\{\mu - 2\sigma < X < \mu + 2\sigma\} = P\{0 < X < 8\} = P(1 \le X \le 7) \qquad [f(x) = e^{-4} 4^x / x!]$$

$$= \sum_{x=1}^{7} f(x) = \sum_{x=0}^{7} f(x) - e^{-4} = 0.949 - 0.018 = 0.931.$$

9*. $M_X(t) = (q + pe^t)^n = [1 + p(e^t - 1)]^n = \left[\dfrac{1 + \lambda(e^t - 1)}{n}\right]^n$

where we used $np = \lambda$ (fixed). Now by Euler's Limit

$$\lim_{n\to\infty} M_X(t) = \lim_{n\to\infty}\left[1 + \frac{\lambda(e^t - 1)}{n}\right]^n = e^{\lambda(e^t - 1)}. \qquad \left[\lim_{n\to\infty}\left(1 + \frac{a}{n}\right)^{bn+c} = e^{ab}\right]$$

10*. Let X and Y denote the number of deaths due to road accidents and due to other causes. Then we are given that X is Pois (2) and Y is Pois (6). By additive property $Z = X + Y$ is Pois (8). Now

$$P(Z \le 10) = \sum_{z=0}^{10} \frac{e^{-8} 8^z}{z!}.$$

Here $\lambda = 8$ which is a positive integer. Hence the maximum probable number of deaths (mode) is 7 or 8.

11*. Let $X_i \sim$ Pois (λ_i) then $S_k = X_1 + \ldots + X_k$ is Pois (λ) where $\lambda = \lambda_1 + \ldots + \lambda_k$. [by reproductive property]. Now

$$\theta = P\{X_1 = r \mid S_k = n\} = P(X_1 = r, S_k = n) / P(S_k = n) = P(X_1 = r, X_2 + \ldots + X_k = n - r) / P(S_k = n)$$

$$= \frac{e^{-\lambda_1} ! \lambda_1^r}{r!} \cdot \frac{e^{-(\lambda - \lambda_1)}(\lambda - \lambda_1)^{n-r}}{(n-r)!} \frac{e^{-\lambda} . \lambda^n}{n!} = \binom{n}{r}\left(\frac{\lambda_1}{\lambda}\right)\left(1 - \frac{\lambda_1}{\lambda}\right)^{n-r} \Rightarrow \quad \theta \sim \text{bin}\,(n, p)$$

where $p = \lambda_1 / \lambda$.

Sec. 13-70. Page 436

1*. If $X = n$ (fixed), then $S \sim$ bin (n, p). Now

$$M(t_1, t_2) = \mathsf{E}\{e^{t_1 S + t_2 D}\} = \mathsf{E}\{e^{(t_1 - t_2)S} e^{t_2 X}\} = \mathsf{E}_X \mathsf{E}\{e^{tS} e^{t_2 X} \mid X = n\} \qquad \text{[by Double-E Rule]}$$

$$= \mathsf{E}_X \{e^{t_2 X} \mathsf{E}(e^{tS} \mid X = n)\} = \mathsf{E}\{e^{t_2 X} .(q + pe^t)^X\}$$

$$= \mathsf{E}\{qe^{t_2} + pe^{t_1})^X\} = \mathsf{E}(\theta^X) \qquad \text{[p.g.f. of Pois } (\lambda)]$$

$$= e^{\lambda(\theta - 1)} = \exp\{\lambda(qe^{t_2} + pe^{t_1}) - \lambda\,(p + q)\}\ .$$

$$= \exp\{\lambda\, p(e^{t_1} - 1)\}.\exp\{\lambda q\,(e^{t_2 - 1})\}$$

Thus, S Pois (λp) and $D \sim$ Pois (λq) are independent distributed.

Remark. The result is false if X is replaced by a fixed number k.

2*. Recall that, if $X \sim$ Pois (λ), then first four moments are

$$\mathsf{E}(X) = \lambda = \text{Var}(X) = \mu_3(X), \qquad \mu_4(X) = \lambda + 3\lambda^2. \qquad \ldots(1)$$

Now, Taylor's expansion about the point λ, for the differentiable function φ is

$$\varphi(x) = \varphi(\lambda) + (x - \lambda)\,\varphi'(\lambda) + [(x - \lambda)^2 / 2!]\,\varphi''(\lambda) + [(x - \lambda)^3 / 3!]\,\varphi'''(\lambda) + [(x - \lambda)^4 / 4!]\,\varphi''''(\lambda) + \ldots$$

Let $\varphi(\lambda) = \lambda^k$, then $\varphi'(\lambda) = k\lambda^{k-1}, \varphi''(\lambda) = k^{(2)}\,\lambda^{k-2}, \varphi'''(\lambda) = \lambda^{(3)}\,\lambda^{k-3}, \ldots$ ...(2)

It follows that Taylor's expansion of the function $\varphi(X) = X^k$ about λ is

$$X^k = \lambda^k + (X - \lambda)\,\varphi'(\lambda) + [(X - \lambda)^2 / 2!]\,\varphi''(\lambda) + [(X - \lambda)^3 / 3!]\,\varphi'''(\lambda) + [(X - \lambda)^4 / 4!]\,\varphi''''(\lambda) + \ldots \quad \ldots(3)$$

Taking expected values of both sides and using (1) and (2) we get

$$\mathsf{E}(X^k) = \lambda^k + \frac{k^{(2)}\,\lambda^{k-1}}{2} + \frac{k^{(3)}\,\lambda^{k-2}}{6} + \frac{(\lambda + 3\lambda^2)\,\lambda^{k-4}\,k^{(4)}}{24} + \ldots \ [\mathsf{E}\,(X - \mu) \equiv 0].$$

$$= \lambda^k + \frac{k(k-1)}{2}\,\lambda^{k-1} + \frac{k(k-1)\,(k-2)\,(3k-5)}{24}\,\lambda^{k-2} + \frac{k(k-1)\,(k-2)\,(k-3)}{24}\,\lambda^{k-3} + \ldots \quad \ldots(4)$$

We replace k by $2k$, $3k$, ... to get higher moments of (X^k). Thus

$$\mathsf{E}(X^{2k}) = \mathsf{E}(Y^2) = \lambda^{2k} + k(2k-1)\,\lambda^{2k-1} + \frac{k(2k-1)\,(k-1)\,(6k-5)}{6}\,\lambda^{2k-2} + \ldots \qquad \ldots(5)$$

The mean value of Y is $\mathsf{E}(Y)$ and is given by (4). Note that

$$\begin{aligned}\mathsf{E}^2(Y) &= \lambda^{2k} + k^{(2)}\lambda^{2k-1} + \tfrac{1}{4}k^2(k-1)^2\lambda^{2k-1} + \tfrac{1}{12}k^{(3)}(3\lambda-5)\lambda^{2k-2} + \dots \\ &= \lambda^{2k} + k(k-1)\lambda^{2k-1} + \tfrac{1}{6}k(k-1)(3k^2-7k+5)\lambda^{2k-2} + \dots\end{aligned}$$

The variance of Y is obtained from $\mu_2(Y) = \mathsf{E}(Y^2) - E^2(Y)$.

$$\begin{aligned}\mu_2(Y) &= \{k(2k-1)\lambda^{2k-1} + \tfrac{1}{6}k(k-1)(2k-1)(6k-5)\lambda^{2k-2} + \dots\} \\ &\qquad - \{k(k-1)\lambda^{2k-1} + \tfrac{1}{6}k(k-1)(3k^2-7k+5)\lambda^{2k+2} + \dots\} \\ &= \lambda^{2k}\left\{\frac{k^2}{\lambda} + \frac{k(k-1)}{6\lambda^2}(9k^2-9k) + \dots\right\} = \lambda^{2k}\left\{\frac{k^2}{\lambda} + \left[\frac{3k^2(k-1)^2}{2\lambda^2}\right] + \dots\right\} \qquad \dots(6)\end{aligned}$$

Putting $k = 1/2$ in (4) and (6), we obtain

$$\begin{aligned}\mathsf{E}(\sqrt{X}) &= \sqrt{\lambda}\,\{1 - (1/8\lambda) - (7/128\,\lambda^2) + \dots\} \\ \mathrm{Var}(\sqrt{X}) &= (1/4) + (3/32\,\lambda) + \dots\end{aligned}$$

Note. The evaluations of $\mathsf{E}(Y^3)$, $\mathsf{E}(Y^4)$, ... are, though straight forward, but time consuming.

3*. *Zero truncated Poisson Distributed.* Since X is Pois (λ), $f(x) = e^{-\lambda}\lambda^x / x!$, $x = 0, 1, 2, \dots$

$$P\{X = x \mid X > 0\} = \frac{f(x)}{P(X>0)} = \frac{e^{-\lambda}\lambda^x / x!}{1-e^{-\lambda}},\ 1 \le x < \infty \qquad [\therefore P(X > 0) = 1 - P(X = 0)]$$

$$\mathsf{E}\{X \mid X > 0\} = \sum_{x=1}^{\infty} \frac{x\,.\,f(x)}{P(X>0)} = \left[\sum_{x=0}^{\infty} x f(x)\right]\Big/(1-e^{-\lambda}) = \frac{\lambda}{1-e^{-\lambda}}$$

As $\qquad P(X > 0, Y > 0) = P(X > 0)\,.\,P(Y > 0) = (1 - e^{-\lambda})\,.\,(1 - e^{-\lambda}) = K^{-1}$ (say)

$$\begin{aligned}\therefore\quad p &\equiv P\{X+Y=n \mid X>0, Y>0\} = K\,P\{X+Y=n, X>0, Y>0\} = K\Sigma P\{X=r, Y=n-r\} \\ &= K\,\Sigma\,P(X=r)\,.\,P(Y=n-r) \qquad [\because \text{ Indep. } (X, Y)] \\ &= K\sum_{r=1}^{n-1}\frac{e^{-\lambda}\lambda^r}{r!}\cdot\frac{e^{-\lambda}\lambda^{n-r}}{(n-r)!} = K\frac{e^{-2\lambda}\lambda^n}{n!}\left[\sum_{r=1}^{n-1}\binom{n}{r}\right] \\ &= \frac{K e^{-2\lambda}\lambda^n}{n!}\left\{\sum_{r=0}^{n-1}\binom{n}{r} - \binom{n}{0} - \binom{n}{n}\right\} = \frac{K e^{-2\lambda}\lambda^n}{n!}\,.\,(2^n - 2) \\ &= \frac{e^{-2\lambda}(2\lambda)^n}{n!}\cdot\frac{1-\left(\frac{1}{2}\right)^{n-1}}{(1-e^{-\lambda})^2}.\end{aligned}$$

4*. Let $P(X = k) = p_k$, then

$$\begin{aligned}\mathsf{E}(t^X) &= p_0 + p_1 t + \dots + p_a t^a + p_{a+1}t^{a+1} + \dots \ge p_0 + p_1 t + \dots + p_a t^a \\ &= t^a\{p_0 t^{-a} + p_1 t^{-a+1} + \dots + p_a\} \ge t^a(p_0 + p_1 + \dots + p_a), \ [\text{for } t^{-k} > 1] \\ &= t^a P\{X \le a\}. \text{[This settles (1)]}\end{aligned}$$

When X is Pois (λ), then $G(t) = \mathrm{E}(t^X) = \exp[\lambda(t-1)]$. ...(2)

From (1) and (2) we get

$$P\{X \le \lambda/2\} \le t^{-(\lambda/2)}\, e^{\lambda(t-1)}, \quad 0 \le t \le 1 \qquad \text{....(3)}$$

Let $z = t^{-(\lambda/2)}\, e^{\lambda(t-1)}, u = \ln z = \frac{1}{2}\lambda \ln t + \lambda(t-1)$.

Differentiating we get ($u' = du/dt$)

$$u' = \lambda - \tfrac{1}{2}\lambda t^{-1},\ u'' = \tfrac{1}{2}\lambda/t^2,\ u' = 0 \Rightarrow t = 1/2 \text{ and } u'' > 0 \text{ at } t = 1/2$$

Hence $\min \ln z = \min u = \frac{1}{2}\ln(2/e) \Rightarrow z = (2/e)^{\lambda/2}$

Since (3) is true $\forall\, t \in [0, 1]$, it must be true at $t = 1/2$. Hence $P(X \le \lambda/2) \le (2/e)^{(\lambda/2)}$.

5*. Note that $P(Z = 1) = P(Z \ge 1) - P(Z \ge 2)$. ...(1)

Now,

$$\begin{aligned}
f(k) &= P(X = k) = e^{-1}/k!, \quad k = 0, 1, 2, \ldots \\
P(Z \ge k) &= P(X \ge k, Y \ge k) = P(X \ge k).\, P(Y \ge k) = [P(X \ge k)]^2. \\
P(Z \ge 1) &= [P(X \ge 1)]^2 = [1 - f(0)]^2 = (1 - e^{-1})^2. \\
P(Z \ge 2) &= [P(X \ge 2)]^2 = [1 - f(0) - f(1)]^2 = (1 - e^{-1} - e^{-1})^2.
\end{aligned}$$

Substituting from above in (1) we get

$$P(Z = 1) = (1 - 2e^{-1} + e^{-2}) - (1 - 4e^{-1} + 4e^{-2}) = (2e - 3)/e^2.$$

6*.

$$\begin{aligned}
G_X(s, t) &= e^{a(s-1)} \Rightarrow X \sim \text{Pois}(a); \text{ so } \mathrm{E}(X) = a = \mathrm{Var}(X). \\
G_Y(1, t) &= e^{b(t-1)} \Rightarrow Y \sim \text{Pois}(b); \text{ so } \mathrm{E}(Y) = b = \mathrm{Var}(Y).
\end{aligned}$$

Let $W = a(s-1) + b(t-1) + c(s-1)(t-1)$; then $G(s, t) \exp W$.

$(\partial G/\partial s) = [a + c(t-1)]\exp(W); (\partial^2 G/\partial t\, \partial s) = [c + [a + c(t-1)][b + c(s-1)]]\exp(W)$.

Put $s = t = 1$, to get $(\partial^2 G(1,1)/\partial s\, \partial t) = \mathrm{E}(X, Y) = c + ab$

$$\mathrm{Cov}(X, Y) = \mathrm{E}(XY) - \mathrm{E}(X)\mathrm{E}(Y) = c, \quad \rho(X\, Y) = c/\sqrt{ab}$$

Note. If $a = b = c = 1$, then $G(s, t) = e^{st} - 1 = G(t, s)$, so X and Y have same p.m.f. as under :

$$G(s, t) = e^{-1}\sum_{r=0}^{\infty}\frac{(st)^r}{r!} \Rightarrow P(X = r, Y = r) = e^{-1}/r!, \quad r = 0, 1, 2, \ldots$$

Chapter 14 : *Geometric Distribution. Negative Binomial Distribution*

Sec. 14-23. Page 442

1*. Here $T_j \sim$ geom (p_j), $j = 1, 2,$; and by independence

$$f(x, y) = f_1(x).\, f_2(y) = (p_1 q_1^{x-1})(p_2 q_2^{y-1}),\ 1 \le x, y \le \infty.$$

$$P\{T_2 > T_1\} = \Sigma P\{T_1 = x, T_2 \ge x+1\} = (p_1\, p_2) \sum_{x=1}^{\infty} q_1^{x-1} \sum_{y=x+1}^{\infty} q_2^{y-1}$$

$$= p_1 p_2 \sum_{x=1}^{\infty} q_1^{x-1} \cdot \left(\frac{q_2^x}{1-q_2}\right) = q_2 p_1 \sum_{x=1}^{\infty} (q_1 q_2)^{x-1} = q_2\, p_1 \cdot \frac{1}{1-q_1 q_2}.$$

By symmetry, $P\{T_1 > T_2\} = p_2 q_1 / (1 - q_1 q_2)$.

Note. $P\{T_1 > T_2\} + P\{T_2 > T_1\} + P\{T_1 = T_2) = 1.$

$$\therefore \quad P(T_1 = T_2) = 1 - \{(p_1 q_2 + p_2 q_1) / (1 - q_1 q_2) = p_1\, p_2 / (1 - q_1 q_2).$$

2*. Let X and Y be the number of shots fired by A and by B respectively. Since X and Y are independent geometric variates

$$P(x_i, y_i) = P(x_i) . P(y_i) = \left(\frac{2}{5}\right)^{x-1} \frac{3}{5} \cdot \left(\frac{2}{7}\right)^{y-1} \frac{5}{7} : (x, y = 1, 2, 3, \ldots)$$

If p is the required probability, then

$$p = P(Y > X) = \sum_{x=1}^{\infty} P(x, y > x) = \sum_{x=1}^{\infty} \sum_{y=x+1}^{\infty} \left(\frac{2}{5}\right)^{x-1} \frac{3}{5} \cdot \left(\frac{2}{7}\right)^{y-1} \frac{5}{7}$$

$$= \sum_{x=1}^{\infty} \left(\frac{2}{5}\right)^{x-1} \cdot \frac{3}{5} \cdot \left(\frac{2}{7}\right)^{x} = \frac{3}{2} \sum_{x=1}^{\infty} \left(\frac{4}{35}\right)^{x} = \frac{3}{2} \cdot \frac{4}{31} = \frac{6}{31}.$$

Note.
$$\sum_{y>x} \left(\frac{2}{7}\right)^{y-1} \frac{5}{7} = \sum_{r=1}^{\infty} \left(\frac{2}{7}\right)^{x+r} \frac{5}{7} = \left(\frac{2}{7}\right)^{x}.$$

3*. Each of the 100 sampled items has probability 0.03 of being defective. The number of defectives in the sample, X, is consequently a binomial variate, and we approximate it by Poisson law. Since $\lambda = np = 100 \times 0.03 = 3$, we have,

$$p = P(X \ge 3) = 1 - P(X \le 2) = 1 - 0.6763 = 0.3237, \text{ (using tables)}.$$

Each of the samples has thus a probability 0.324 of stopping the machine. The number of samples taken between the successive stoppages is, consequently, a geometric variate with mean $1/p = (1/0.324) = 3.1$. Thus the average time interval between Successive adjustments is 3.1 hrs.

4*. Here
$$f(x) = (1/2)^x = pq^{x-1}, \; (p = q = 1/2), \quad 1 \le x \le \infty.$$

$$\therefore \quad E(X) = 1/p = 2; \; \sigma^2 = \text{Var}(X) = q/p^2 = 2. \quad [X \sim \text{gem}(p)]$$

Chebyshev's inequality : $P\{|X - \mu| \le c\} > 1 - (\sigma^2 / c^2) \Rightarrow P\{|X - 2| \le 2\} > 1 - \frac{1}{2} = \frac{1}{2}$.

Actual probability p_1 is given by

$$p_1 = P\{|X - 2| \le 2\} = P\{0 \le X \le 4\} = \Sigma P(X = k), \quad k = 1, 2, 3, 4.$$

Note that $X = 0$ is not defined by given p.d.f. so $f(x) = 0$ at $x = 0$. Now

$$p_1 = \left(\tfrac{1}{2}\right) + \left(\tfrac{1}{2}\right)^2 + \left(\tfrac{1}{2}\right)^3 + \left(\tfrac{1}{2}\right)^4 = 15/16.$$

5*. Notice that $Z = \max(X, Y) \le m \Rightarrow \{(X \le m) \cap (Y \le m)\}$. Thus using indep. & c.d.f.

$$P[\max(X, Y) \le 1] = P(X \le m).P(Y \le m) = (1-q^m).(1-q^m) = (1-q^m)^2.$$

$$\therefore \quad P\{Z = m\} = P\{Z \le m\} - P\{Z \le m-1\} = (1 - 2q^m + q^{2m}) - (1 - 2q^{m-1} + q^{2m-2})$$

$$= 2q^{m-1}(1-q) - q^{2m-2}(1-q^2) = 2pq^{m-1} - p(1+q)q^{2m-2}. \quad [\because q = 1-p].$$

Sec. 14-51. Page 446

1*. Recall that for geom (p), $M(t) = p(1 - qe^t)^{-1}$. Comparing with the given m.g.f. $= (1/5)[1 - (4/5)e^t]^{-1}$ we see that X is geom (1/5). Hence $P(X = x) = q^x p$, $x = 0, 1, 2, \ldots$

$$\therefore \quad P\{X = 5 \text{ or } 6\} = P(X=5) + P(X=6) = pq^5 + pq^6 = pq^5(1+q) = 9 \times 4^5/5^7.$$

Note. If m.g.f. is $e^t(5 - 4e^t)^{-1}$, then $X \sim$ geom (1/5).

2*. Let X and Y be the number of throws required by A and B (separately) to obtain their objectives; then X is geom (p_1) and Y is geom (p_2), where $p_1 = 1/6$, $p_2 = 2/6 = 1/3$. Hence $f(x) = q_1^{x-1}p_1$, $g(y) = q_2^{y-1}p_2$; $x, y = 1, 2, \ldots$

$$p = P(Z = k) = P(X + Y = k) = \sum_{j=1}^{k} P(Y = j, X = k - j) = \sum_{j=1}^{k} P(Y = j)\, P(X = k - j)$$

$$= \sum_j (q_2^{j-1}p_2)(p_1 q_1^{k-j-1}) = p_1 p_2 q_1^{k-2} \sum_{j=1}^{k-1} (q_2/q_1)^{j-1}. \quad \left(\frac{q_2}{q_1} = \frac{4}{5}\right)$$

$$= \frac{1}{18}\left(\frac{5}{6}\right)^{k-2}\left[\frac{1-(4/5)^{k-1}}{1-(4/5)}\right] = \frac{5}{18}\left(\frac{5}{6}\right)^{k-2} - \frac{2}{9}\left(\frac{2}{3}\right)^{k-2}.$$

Method of generating functions. $G_X(t) = p_1 t^1(1 - q_1 t)$, $G_Y(t) = p_2 t^1(1 - q_2 t)$; $Z = X + Y$.
$G_Z(t) = G_X(t) \,.\, G_Y(t) = p_1 p_2 t^2/(1 - q_1 t)(1 - q_2 t) = t^2/18\,(1 - 2t/3)(1 - 5t/6)$

$$= \frac{5}{18}\frac{t^2}{1-(5t/6)} - \frac{2}{9}\frac{t^2}{1-(2t/3)} = \frac{5}{18}t^2 \sum_{r=0}^{\infty}\left(\frac{5t}{6}\right)^r - \frac{2}{9}t^2 \sum_{r=0}^{\infty}\left(\frac{2t}{3}\right)^r \quad \ldots(1)$$

$$\therefore \quad P(Z = k) = \text{coeff. of } t^k \text{ in } (1) = (5/18)(5/6)^{k-2} - (2/9)(2/3)^{k-2}.$$

3*. Note that $P(X \le x) = \sum_{k=0}^{x} pq^k = p(1 - q^{x+1})/(1-q) = 1 - q^{x+1}$.

(a) $P\{Z \le x\} = P(\max(X,Y) \le z\} = P\{X \le z, Y \le z\} = P(X \le z)\,P(Y \le z) = (1 - q^{z+1})^2$.

$\therefore \quad P\{Z = n\} = P(Z \le n) - P(Z \le n-1) = (1-q^{n+1})^2 - (1-q^n)^2 = pq^n(2 - q^n - q^{n+1}) \ldots(1)$

(b) To find $P\{Z = n, X = k\}$, note that if $n < k$, then $\{Z = n\} \cap \{X = k\} = \varnothing$ so $P(n, k) = 0$. Thus, we need to consider only $n = k$ or $n > k$.

For $n = k$, we have, writing $P(X = a, Y = b) = P(a, b)$

$$P(X = k, Z = k) = P\{X = k, \max(X, Y) = k\} = P\{(k, 0) \uplus (k, 1) \uplus \ldots \uplus (k, k)]$$

$$= P(k, 0) + P(k, 1) + \ldots + P(k, k)$$

$$= P(X = k)\sum_{j=0}^{k} P(Y = j) = (pq^k)\sum_{j=0}^{k} pq^j = pq^k(1 - q^{k+1}). \quad [X \;\&\; Y \text{ are Indep.}]$$

For $n > k$, we have

$$f(k, n) = P\{X = k, Z = n\} = P\{Y = n, X = k\} = P(Y = n)\, P(X = k) = (pq^n)(pq^k) = p^2 q^{n+k}.$$

$$\therefore \quad P(X = k, Z = n) = \begin{cases} 0, & n < k \\ pq^k (1 - q^{k+1}), & n = k = 0, 1, 2 \dots \quad \dots(2) \\ p^2 q^{k+n}, & n > k = 0, 1, 2, \dots \quad \dots(3) \end{cases}$$

Observe that $\sum P(X = k, Z = n) = \sum f(k, n) = 1$; as $\sum f(k, n)$ is

$$\sum_{k=0}^{\infty} pq^k (1 - q^{k+1}) - \sum_{k=0}^{\infty} \left[\sum_{n=k+1}^{\infty} p^2 q^{k+n} \right] = \sum_{k=0}^{\infty} pq^k (1 - q^{k+1}) - \sum_{k=0}^{\infty} (pq^{2k+1}) = \sum_{k=0}^{\infty} pq^k = 1.$$

(c) $$P(X = k \mid Z = n) = \frac{f(k,n)}{P(Z = n)} = \begin{cases} (1 - q^{k+1}) / (2 - q^{k+1} - q^k) \\ pq^k / (2 - q^{n+1} - q^n, \text{ if } n > k. \\ 0, \text{ if } n < k \end{cases}$$ [use $n = k$ in (2) & (1)]

(d) $$P(Z = n \mid X = k) = \frac{f(k,n)}{P(X = k)} = \begin{cases} pq^k (1 - q^{k+1}) / pq^k = 1 - q^{k+1} & (n = k) \\ p^2 q^{n+k} / pq^k = pq^n, & \text{if } (n > k = 0, 1, 2, \dots.) \\ 0, & \text{if } (n < k) \end{cases}$$

4*. We define random variables $X_1, X_2, \dots, X_n$ as follows :

X_1 = No. of balls that must be distributed for any one cell (say the first cell) be occupied. [Obviously $X_1 = 1$ because the first ball automatically occupies one cell.]

X_2 = additional No. of balls that must be distributed to occupy any two cells.

X_3 = additional No. of balls that must be distributed to occupy any three cells, etc.

Once one cell is occupied, as successive balls are distributed, failures occur when these balls enter the already occupied cell. A success occurs when one of the available $N - 1$ unoccupied cells is finally entered. The Prob. of success is $p_2 = (N - 1)/N$. Thus X_2 is geom (p_2) where $p_2 = (N - 1)/N$; X_3 is geom (p_3) where $p_3 = (N - 2)/N$, ... and finally X_k is geom (p_k) where

$p_k = [N - (k - 1)]/N$, etc. We observe that X_i are indep. geometric variates; hence if $X = X_1 + X_2 + \dots. X_N$, then

$$G(t : X) = G(t : X_1) \,.\, G(t : X_2) \dots G(t : X_N).$$

Substitute the values of $G(t : X_j)$:

$$G(t : X) = \frac{p_1 t}{1 - q_1 t} \cdot \frac{p_2 t}{1 - q_2 t} \cdots \frac{p_N t}{1 - q_N t} \qquad (q = 1 - p;\ p_1 = 1)$$

$$= t^N \left(\frac{N-1}{N} \cdot \frac{1}{1 - t/N} \right) \left(\frac{N-2}{N} \cdot \frac{1}{1 - 2t/N} \right) \cdots \left(\frac{1}{N} \cdot \frac{1}{1 - (N-1)t/N} \right)$$

$$= N!\theta^N (1 - \theta)^{-1} (1 - 2\theta)^{-1} \dots [1 - (N - 1)\theta]^{-1} \,.\, [t/N = \theta]$$

Sec. 14-72. Page 454

1*. The number of trials X on which kth success occurs has the Neg-bin density :

$$f(x) = \binom{x-1}{k-1} p^k q^{x-k}, x = k, k+1, \ldots$$

Here $\quad p = 1/2 = q, x = 10, k = 5.$

$$\therefore \quad P = \binom{9}{4}\left(\frac{1}{2}\right)^5 \left(\frac{1}{2}\right)^5 = \frac{9!}{4!\ 5!}\left(\frac{1}{2}\right)^{10} = \frac{63}{(2)^9}.$$

2*. To get 2 defectives, we have to perform at least two trials (drawing one item at a time). The probability of a success is 0.03 for every trial. The situation admits negative binomial model; hence the required probability is

$$P = \{X \geq 5\} = \sum_{x=5}^{\infty} \binom{x-1}{2-1} p^2 q^{x-2} = 1 - \sum_{x=2}^{4} \binom{x}{1} (0.03)^2 (0.97)^{x-2}$$

$$= 1 - 0.0009\ (2 + 2.91 + 3.7636) = 1 - 0.0009 \times 8.6736 = 0.9922.$$

Comments. *Neg-bin as difference of two sums of Binomial Probabilities*

If X is the number of failures preceding the kth success, $[X \sim \text{NB}(k, p)]$ then

$$p(x) = \binom{k+x-1}{x} p^k q^x, \quad x = 0, 1, 2, \ldots \qquad \ldots(1)$$

$P(X > x) = P\{$more than $\{x + k)$ trials are required to obtain k successes$\}$

$$= P\{(x = k) \text{ trials produced fewer than } k \text{ successes}\} = \sum_{j=0}^{k-1} \binom{x+k}{j} p^j q^{x+k-j}$$

$$p(x) = P\{X > x-1\} - P(X > x\} = \sum_{j=0}^{k-1} \binom{x+k-1}{j} p^j q^{x-1+k-j} - \sum_{j=0}^{k-1} \binom{x+k}{j} p^j q^{x+k-j}.$$

Sec. 14-94. Page 459

1*. We know that $\mu = \text{E}(X) = kq/p$, and so, by definition

$$\mu_r = \sum_{x=0}^{\infty} \binom{k+x-i}{x} q^x p^k \left(x - \frac{kq}{p}\right)^r. \qquad \ldots(1)$$

We assume that the infinite series (1) is uniformly convergent, so that term by term differentiation is legitimate. We differentiate (1) w.r.t. "q" and noting that $(\partial p / \partial q) = -1$, we obtain

$$\frac{\partial \mu_r}{\partial q} = \sum_{x=0}^{\infty} \binom{k+r-1}{x} \frac{\partial}{\partial q}\left[q^x p^k \left(x - \frac{kq}{p}\right)^r\right]$$

$$= \sum_{x=0}^{\infty} \binom{k+r-1}{x}\left[(xq^{x-1}p^k - kq^x p^{k-1}\left(x - \frac{kq}{p}\right)^r - rkq^x p^k\left(x - \frac{kq}{p}\right)^{r-1} \cdot \left(\frac{p+q}{p^2}\right)\right]$$

$$= \sum_{x=0}^{\infty} \left\{ \binom{k+x-1}{x} q^x p^k \right\} \left[\left(\frac{x}{q} - \frac{k}{p} \right) (x-\mu)^r - \frac{kr}{p^2} (x-\mu)^{r-1} \right]$$

$$= \sum_{x=0}^{\infty} f(x) \left[\frac{1}{q} (x-\mu)^{r+1} - \frac{kr}{p^2} (x-\mu)^{r-1} \right]$$

$$= \frac{1}{q} \mathrm{E}(X-\mu)^{r+1} - \frac{kr}{p^2} \mathrm{E}(X-\mu)^{r-1} = \frac{\mu_{r+1}}{q} - \frac{rk}{p^2} \mu_{r-1}.$$

Thus by transfer : $\mu_{r+1} = q \left(\dfrac{\partial \mu_r}{\partial p} + \dfrac{rk}{p^2} \mu_{r-1} \right).$

2*. By definition of Survival function $P\{X \geq m\} = G$ (say), we get

$$G = P\{X \geq m\} = \sum_{r=m}^{\infty} \binom{-k}{r} Q^{-k-r} (-P)^r = \sum_{r=m}^{\infty} \binom{k+r-1}{r} Q^{-k-r} P^r.$$

This series is uniformly convergent, its sum being less than unity. Term by term differentiation leads to

$$\frac{dG}{dP} = \sum_{r=m}^{\infty} \binom{k+r-1}{r} [r P^{r-1} Q^{-(k+r)} - (k+r) P^r Q^{-(k+r+1)}]$$

$$= \sum_{r=m}^{\infty} \left\{ r \binom{k+r-1}{r} P^{r-1} Q^{-(k+r)} - (r+1) \binom{k+r}{1+r} P^r Q^{-(k+r+1)} \right\}$$

$$= \sum_{r=m}^{\infty} \{\phi(r) - \phi(r+1)\} = \phi(m), \text{ as } \phi(\infty) = 0, \left[\phi(r) = r \binom{k+r-1}{r} P^{r-1} Q^{-(k+r)} \right].$$

$$\therefore \quad \frac{dG}{dP} = m \binom{k+m-1}{r} P^{m-1} Q^{-(k+m)} = \frac{P^{m-1} Q^{-(k+m)}}{B(m,k)}$$

Integrating this result w.r.t. "P" in the range $(0, a)$ we get

$$G = \frac{1}{B(m,k)} \int_0^a x^{m-1} (1+x)^{-(k+m)} \, dx \qquad [\because Q = 1 + P]$$

Replacing a by P, reverting to original notation we get

$$P\{X \geq m\} = \frac{1}{B(m,k)} \int_0^P \frac{x^{m-1} \, dx}{(1+x)^{m+k}}.$$

Note. $\phi(\infty) = \lim\limits_{m \to \infty} \dfrac{1}{B(m,k)} \left(\dfrac{P}{Q} \right)^m \cdot \dfrac{1}{(PQ)^k} = \lim\limits_{m \to \infty} \dfrac{(PQ^k)^{-1}}{\Gamma(k)} \dfrac{\Gamma(m+k)}{\Gamma(m)} q^m, \quad \left[q = \dfrac{P}{Q} < 1 \right]$

As $\dfrac{\Gamma(m+k)}{\Gamma(m)} \approx m^k, \lim\limits_{m \to \infty} \dfrac{q^m}{m^{-k}} = 0.$ Hence $\phi(\infty) = 0.$ [Vide § 0-11 (i)]

3*. Recall : $(1-t)^{-n} = \sum\limits_{x=0}^{\infty} \dbinom{n+x-1}{x} t^x = \sum\limits_{x=0}^{\infty} \dfrac{n(n+1)\ldots(n+x-1)}{x!} t^x = \sum\limits_{x=0}^{\infty} \dfrac{n^{[x]} t^x}{x!}$ ...(1)

Also $\Gamma(k+x)/\Gamma(k) = k^{[x]} = k(k+1)\ldots(k+x-1).$ [Reverse Factorial powers]

Here $f(x) = \dfrac{\Gamma(k+x)}{\Gamma(k)\,x!}\left(\dfrac{a}{a+t}\right)^k\left(\dfrac{t}{a+t}\right)^x = \dfrac{k^{[x]}}{x!}\,p^k\,q^x, 0 \le x < \infty,\ \ f(0) = \left(\dfrac{a}{a+t}\right)^k = (p)^k.$ ...(2)

Now $\sum\limits_{x=0}^{\infty} f(x) = 1 \Rightarrow \sum\limits_{x=1}^{\infty} f(x) = 1 - f(0) \Rightarrow \sum\limits_{x=1}^{\infty} f(x)/[1-f(0)] = 1 = \lambda \sum\limits_{x=1}^{\infty} f(x), \lambda = [1-f(0)]^{-1}.$

The zero-truncated density is thus

$$f(x) = \lambda k^{[x]} p^k q^x / x!, \quad 1 \le x < \infty, \quad [x! = x^{(s)}\{(x-s)!\}]$$

$$\therefore \quad \mathsf{E}(X^{(s)}) = \lambda p^k . \sum_{x=1}^{\infty} \frac{x^{(s)}\,q^x k^{[x]}}{x!} = \lambda p^k \sum_{x=s}^{\infty} \frac{q^x\,k^{[s]}}{(x-s)!} = \lambda p^k q^s \sum_{y=0}^{\infty} \frac{q^y k^{[s+y]}}{y!}. \quad [y = x - s]$$

$$= \lambda p^k q^s .k(k+1)\ldots(k+s-1) \sum_{y=0}^{\infty} \frac{q^y (k+s)^{[y]}}{y!}. \quad [k^{[y+s]} = k(k+1)\ldots(k+s-1)(k+s)^{[y]}]$$

$$= \lambda p^k q^s\, k(k+1)\ldots(k+s-1)(1-q)^{-(k+s)} \quad \text{[by (1)]}$$

$$= \lambda (q/p)^s\, k(k+1)\ldots(k+s-1) = \lambda (t/a)^s .k(k+1)\ldots(k+s-1). \quad [q/s = t/a = \theta]$$

$$\therefore \mathsf{E}(X) = \lambda k (t/a), \mathsf{E}[X(X-1)] = \lambda\, k(k+1)\theta^2, \mathsf{E}[X(X-1)(X-2)] = \lambda k(k+1)(k+2)\theta^3$$

$$\mathsf{E}(X^2) = \lambda k\theta\,[1+(k+1)\theta], \quad \mathrm{Var}(X) = \lambda\, k\theta\{(k+1)\theta + 1 - \lambda\, k\theta\}.$$

$$\mathsf{E}(X^3) = \lambda k(k+1)(k+2)\theta^3 + 3\lambda k\theta\{1+(k+1)\theta\} - 2\lambda\, k\theta.$$

$$A \equiv \mu_2'/\mu_1' = 1 + (k+1)\theta; B \equiv \mu_3'/\mu_1' = 1 + 3(k+1)\theta + (k+1)(k+2)\theta^2$$

$$\therefore \quad A - 1 = (k+1)\theta, \ (B-1) - 3(A-1) = (A-1)(k+2)\theta.$$

Dividing : $\dfrac{k+2}{k+1} = \dfrac{B-3A+2}{(A-1)^2}$, which yields k and hence $\theta = t/a$ is expressible in terms of μ_r'.

4*. Let $P_r(n)$ represent the probability that the nth trial (T_r) is the rth success (S_r) : Now observe the obvious :

(i) T_1 is a success and there remain $(n-1)$ trials to reach further $(r-1)$ successes ;

(ii) T_1 is a failure and all r successes must occur in $(n-1)$ trials.

Events (i) and (ii) are disjoint, so by Rule of Total probability.

$$P_r(n) = p\,P_{r-1}(n-1) + q\,P_r(n-1), \quad 2 \le r \le n$$

Multiply by t^n and add up to obtain

$$\left\{\sum_{n=r}^{\infty} t^n P_r(n) = p \sum_{n=r}^{\infty} t^n\, P_{r-1}(n-1) + q \sum_{n=r}^{\infty} t^n\, P_r(n-1)\right\} \Rightarrow G_r(t) = pt\,G_{r-1}(t) + qt\,G_r(t).$$

$$\therefore \quad G_r(t) = [pt/(1-qt)]\,G_{r-1}(t) = [pt/(1-qt)]^2\, G_{r-2}(t) = \ldots = [pt/(1-qt)]^{r-1}\, G_1(t)$$

by repeated reduction of r. Now

$$G_1(t) = \mathsf{E}(t^X) = \Sigma t^k (pq^{k-1}) = pt(1 + qt + q^2t^2 + \ldots) = pt/(1-qt).$$

Thus $$G_r(t) = [pt/(1-qt)]^r.$$

Chapter 15 : More Discrete Distributions

Sec. 15-12. Page 463

1*. Since X is $U\{0, 1, ..., N\}$, etc $P(X = r\} = 1/(N + 1) = P\{Y = r\}$, $0 \le r \le N$.

(i) $$P\{X \ge Y\} = \Sigma P\{Y = r, X \ge r\} = \Sigma P(Y = r) P(X \ge r), \quad \text{[indep. } (X, Y)]$$

$$= \frac{1}{(N+1)} \sum_{r=0}^{N} P(X \ge r) = \frac{1}{N+1} \mathsf{E}(X) = \frac{1}{2} \frac{N}{N+1}.$$

Recall that $\mathsf{E}(X) = \Sigma P(X \ge r)$, if r is an integer [vide Chap. 5].

(ii) $$P\{Z \ge z\} = P\{X \ge z, Y \ge z\} = P(X \ge z) P(Y \ge z) = [P(X \ge z)]^2 = [1 - P(X < z)]^2.$$

$$P(X < z) = \sum_{r=0}^{z-1} P(X = r) = \frac{1}{N+1} \{1 + 1 + ... + 1\}_{z\text{-terms}} = \frac{z}{N+1}.$$

$$\therefore \quad P\{Z = z\} = P\{Z \ge z\} - P\{Z \ge z+1\} = \left(1 - \frac{z}{N+1}\right)^2 - \left(1 - \frac{z+1}{N+1}\right)^2$$

$$= [2(N - z) + 1] / (N + 1)^2, \quad z = 0, 1, 2, ..., N.$$

2*. We consider all the $n + 1$ random observations together. One of them has to be the minimum. Each has the same chance of being the smallest. Hence, the probability that the last observation is the smallest is $p = 1/(n + 1)$.

Sec. 15-27. Page 472

1*. If X denotes the number of defective items, then the p.d.f. is

$$f(x) = \binom{M}{x}\binom{N-M}{n-x} \Big/ \binom{N}{n}, \quad x \in \{0, 1, 2, ..., M\}$$

$$\therefore \quad \phi(t) = \mathsf{E}(e^{itX}) = \sum_{x=0}^{m} \binom{M}{x}\binom{N-M}{n-x} e^{itx} \Big/ \binom{N}{n}.$$

Remark. Characteristic function of Hyp-geom distribution is not convenient as a means of finding the moments of this distribution.

2*. Let X be the random variable denoting the number of judges favouring A. Then the p.d.f. of X is

$$P\{X = x\} = \binom{4}{x}\binom{3}{3-x} \Big/ \binom{7}{3}, \quad x = 0, 1, 2, 3.$$

Let $$M = \{X = 3\} \uplus \{X = 2\}, \text{ hence}$$

$$P(M) = P(X = 3) + P(X = 2) = \left[\binom{4}{3}\binom{3}{0} + \binom{4}{2}\binom{3}{1}\right] \Big/ \binom{7}{3} = \frac{(4+18)}{35} = \frac{22}{35}.$$

3*. We shall make appropriate changes in the formula of p.m.f.

$$f(x; N, M, N) = \binom{M}{x}\binom{N-M}{n-n}\Big/\binom{N}{n}, \quad \left\{\binom{n}{r} = \binom{n}{n-r}\right\} \qquad \ldots(4)$$

(i) $x \to n - x$, $M \to N - M$ provides

$$\binom{N-M}{n-x}\binom{N-(N-M)}{n-(n-x)}\Big/\binom{N}{n} = \binom{N-M}{n-x}\binom{M}{x}\Big/\binom{N}{n}.$$

Thus $f(n - x, N, N - M, n) = f(x\,;\, n, N, M)$.

(ii) $x \to M - x$, $n \to N - m$ provides

$$\binom{M}{M-x}\binom{N-M}{(N-n)-(M-x)}\Big/\binom{N}{N-n} = \binom{M}{M-x}\binom{N-M}{N-M-(n-x)}\Big/\binom{N}{N-n}.$$

By (4b), the R.H.S. is simply $f(x\,;\, n, N, M)$; whence (2) is established.

(iii) $x \to N - n - M + x$, $\quad n \to N - n$, $M \to N - M$ provides

$$\binom{N-M}{N-M-(n-x)}\binom{N-(N-M)}{(N-n)-(N-n-M+x)}\Big/\binom{N}{N-n} = \binom{N-M}{N-M-(n-x)}\binom{M}{M-x}\Big/\binom{N}{N-n}$$

By (4b), the R.H.S. is simply $f\,(x;\, n, N, M)$, whence (3) is established.

4*. Using Partition Theorem :

$$p_r = P\{X + Y = r\} = \sum_{k=0}^{r} P\{X = k, Y = r - k\} = \Sigma\, P(X = k)\,.\,P\{Y = r - k \mid X = k\}$$

$$= \sum_{k=0}^{r} \frac{\binom{50}{k}\binom{150}{10-k}}{\binom{200}{10}} \cdot \frac{\binom{50-k}{r-k}\binom{140+k}{10-r+k}}{\binom{190}{10}}.$$

We expand the binomial coefficients to simplify this expression, cancel terms and obtain after *adjusting*

$$p_r = \frac{150!(10!)^2}{200!}\,\frac{50!}{(50-r)!}\,\frac{150!}{(130+r)!} \times \sum_{k=0}^{r} \frac{1}{k!}\,\frac{1}{(r-k)!}\,\frac{1}{(10-r+k)!}\left(\frac{r!}{r!}\right)\left(\frac{(20-r)!}{(20-r)!}\right)$$

$$= \frac{\binom{50}{r}\binom{150}{20-r}(10!)^2}{\binom{200}{20}20!} \sum_{k=0}^{r}\binom{r}{k}\binom{20-r}{10-k} = \binom{200}{20}^{-1}\binom{50}{r}\binom{150}{20-r}\frac{(10!)^2}{20!}\binom{20}{10}$$

$$= \binom{50}{r}\binom{150}{20-r}\Big/\binom{200}{20} \;:\; \text{p.m.f. of H-G } (200, 50\,;\, 20).$$

Sec. 15-36. Page 479

1*. $P\{\text{all faces alike}\} = 6/6^3 = 1/36 = p_1$ [use Bin - distribution]

$$P\{\text{only two faces alike}\} = \frac{3!}{2!\ 1!}\frac{6\times 5}{6^3} = \frac{15}{36} = p_2.$$

The joint distribution of X and Y is trinomial with $n = 10$, $p_1 = 1/36$ and $p_2 = 15/36$, hence

$$f(x, y) = \frac{10!}{x!\,y!\,(10-x-y)!}\left(\frac{1}{36}\right)^x\left(\frac{15}{36}\right)^y\left(1-\frac{1}{36}-\frac{15}{36}\right)^{10-x-y};\ 0 \le x+y \le 10.$$

Since $\mathrm{E}(XY) = n(n-1)\,p_1\,p_2$, [§ 15-35] we get for the problem under consideration

$$\mathrm{E}(24XY) = 24\ .\ 10\ .\ 9\ .\ (1/36)\ (15/36) = 25.$$

2*. $$M(t:Z) = (e^{tz}) = \mathrm{E}(e^{tn-tX-tY}) = e^{tn}\,\mathrm{E}(e^{t_1X+t_2Y})\text{ where } t_1 = -t = t_2.$$

$$M(t) = e^{tn}\,[1-p-q+(p+q)\,e^{-t}]^n = [p+q+(1-p-q)\,e^t]^n.\ [\S\ 15\text{-}53,\ p_1 = p,\ p_2 = q]$$

$$\mathrm{E}(Z) = M'(0) = n(1-p-q);\ \mathrm{E}(Z^2) = M''(0) = n(1-p-q)\,[(n-1)\,(1-p-q)+1]$$

$$\mathrm{Var}\,(Z) = \mathrm{E}(Z^2) - \mathrm{E}^2(Z) = n(p+q)\,(1-p-q).$$

$$\mathrm{Cov}\,(Y, Z) = \mathrm{Cov}(Y, n-X-Y) = -\mathrm{Cov}\,(Y, X) - \mathrm{Var}\,(X) = npq - nq\,(1-q) = -nq(1-p-q).$$

3*. Let $k_1 + \ldots + k_m = r$, so that $k_{m+1} = n - r$. Also $p_{m+1} = 1 - \sum_1^m p_j = 1 - \left(\sum_1^m \lambda_j / n\right)$

The multinomial law can now be written as

$$P = \frac{n!}{(n-r)!\prod(k_j!)}\left(\prod_{j=1}^m p_j^{k_j}\right)\left(1-\frac{\Sigma\lambda_j}{n}\right)^{n-r}$$

Since $$\frac{n!}{(n-r)!} = n\,(n-1)\ldots(n-r+1) = n^r\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\left(1-\frac{r-1}{n}\right)$$

$$\therefore\quad P = \frac{n^r\prod(p_j)^{k_j}}{\prod(k_j!)}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\ldots\left(1-\frac{r-1}{n}\right)\left(1-\frac{\Sigma\lambda_j}{n}\right)^{n-r}$$

$$= \frac{\prod(\lambda_j)^{k_j}}{\prod(k_j!)}\left(1-\frac{\Sigma\lambda_j}{n}\right)^{n-r}.\left(1-\frac{1}{n}\right)\ldots\left(1-\frac{r-1}{n}\right), (r = \Sigma k_i)$$

Now using the limit, $(1 - a/n)^{bn+c} \to c^{ab}$ as $n \to \infty$, the above result reduces to

$$P = e^{-\Sigma\lambda_j}\prod(\lambda_j)^{k_j}\Big/\prod(k_j!),\ j = 1, 2, \ldots, m.$$

4*. For $M(n, p_1, p_2, \ldots, p_k)$, $\sigma_{ij} = -np_i\,p_j\ (i \ne j)$, $\sigma_{ij} = n\,p_i\,q_i$, $(p_i + p_j = 1)$. Hence taking n common out of each row of the det $\{\sigma_{ij}\}$, we get

$$\det[\sigma_{ij}] = \begin{vmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{ik} \\ \sigma_{21} & \sigma_{21} & \cdots & \sigma_{2k} \\ \cdots & \cdots & \cdots & \cdots \\ \sigma_{k1} & \sigma_{k2} & \cdots & \sigma_{kk} \end{vmatrix} = (n)^k \begin{vmatrix} p_1 q_1 & -p_1 p_2 & \cdots & -p_1 p_k \\ -p_2 p_1 & p_2 q_2 & \cdots & -p_2 p_k \\ \cdots & \cdots & \cdots & \cdots \\ -p_k p_1 & -p_k p_2 & \cdots & p_k q_k \end{vmatrix}$$

Let S_j denote the sum of elements in the jth row of the last determinant. Then

$$S_1 = p_1(q_1 - p_2 - p_3 \ldots - p_k) = p_1[1 - \Sigma p_i] = p_i(1-1) = 0 \quad [\Sigma p_i = 1]$$

$$S_2 = p_2(-p_1 + q_2 - p_3 \ldots - p_k) = p_2[1 - \Sigma p_i] = p_2(1-1) = 0$$

$$S_k = p_k(-p_1 - p_2 \ldots - p_{k-1} + q_k) = p_k[1 - \Sigma p_i] = p_k(1-1) = 0.$$

Thus adding col 2, col. 3, ... col. k to col. 1 and using $S_j = 0$ ($j = 1, 2, \ldots, k$) we get

$$[\sigma_{ij}] = \det[\sigma_{ij}] = n^k \begin{vmatrix} 0 & -p_1p_2 & -p_1p_3 & \cdots & -p_1p_k \\ 0 & p_2q_2 & -p_2p_3 & \cdots & -p_2p_k \\ \cdot & \cdots & \cdots & \cdots & \cdots \\ 0 & -p_3p_2 & -p_kq_k & \cdots & p_kq_k \end{vmatrix}$$

The value of this determinant = 0, as col 1 consists of zeros. Hence rank $|\sigma_{ij}| \neq k$. However, the det. of order $(k-1) \times (k-1)$ occurring on the right of col. 1 and below row 1 is non-zero. It follows that rank $|\sigma_{ij}| = k - 1$.

5*. For the Die Events; $P(A) = 4/6$ and $P(B) = 2/6$. By Multi-Stage Rule

$$P(X = k) = P(A)\,P(X = k \mid A) + P(B)\,P(X = k \mid B) \Rightarrow p_k = \tfrac{2}{3} P(X \mid A) + \tfrac{1}{3} P(X \mid B)$$

$$p_1 = \frac{2}{3}\binom{30}{10, 10, 10}(0.5)^{10}(0.3)^{10}(0.2)^{10} + \frac{1}{3}\binom{50}{10}\binom{30}{10}\binom{20}{10}\Big/\binom{100}{30} \cdot \text{[trinomial and biv H-G]}$$

$$p_2 = \frac{2}{3}\binom{30}{12}(0.3)^{12}(0.7)^{18} + 0. \qquad p_3 = 1 - P(X=1) - P(X=2) = 1 - p_1 - p_2.$$

We have used trinomial and bivariate H-G to obtain p_1, binomial $[P(X \mid A) = 0]$ to get p_2.

Now : $E(X) = 1P(X=1) + 2P(X=2) + 3P(X=3) = 3 - 2p_1 - p_2$

$$= 3 - 2\left[\frac{2}{3} \cdot \frac{30!}{(10!)^3}(0.03)^{10} + \frac{1}{3}\binom{50}{10}\binom{30}{10}\binom{20}{10}\Big/\binom{100}{30}\right] - \frac{2}{3} \cdot \frac{30}{12!18!}(0.3)^{12}(0.7)^{18}.$$

Chapter 16 : Normal (Gaussian) Distribution

Sec. 16-24. Page 489

1*. We normalize the variates X and Y. Thus $(X - 30)/5 = Z$, i.e. $X = 30 + 5Z$; $(Y - 15)/10 = Z$, i.e. $Y = 15 + 10Z$. Thus if Z is $N(0, 1)$, then

$$P\{26 \le X \le 40\} = P(26 \le 30 + 5Z \le 40) = P\{-0.8 \le Z \le 2\} \quad \ldots(i)$$

$$P\{7 \le Y \le 35\} = P(7 \le 15 + 10Z \le 35) = P\{-0.8 \le Z \le 2\} \quad \ldots(ii)$$

From (i) and (ii) the result follows trivially.

2. Here, we substitute Normal variate for a binomial variate $X \sim$ bin $(2000, 1/2)$.

$\mu \quad = np = 1000, \sigma^2 = npq = 500$; let $Z = (X - \mu)/\sigma$ i.e. $X = \mu + \sigma Z = 1000 + \sqrt{500}\, Z$; then

$$p_0 = P\{900 < X < 1100\} = P\{900 < 1000 + \sqrt{500}\, Z < 1100\}$$

$$= P\{-2\sqrt{5} < Z < 2 < 2\sqrt{5}\} = F(2\sqrt{5}) - F(-2\sqrt{5}) = 2F(2\sqrt{5}) - 1. \quad \text{[Draw a Picture]}$$

Note. The use of $N(\mu, \sigma^2)$ for bin (n, p) for large $n(> 30)$ is justified by CLT. [or limiting distributions]. See Chapter 17.

3*. We use *normalized* r.v. $Z = (X - \mu)/\sigma = (X - 30)/5$ i.e. $X = 30 + 5Z$. [Draw Fig.]

$$p_1 = P(26 \le X \le 40) = P(26 \le 30 + 5Z \le 40) = P(-0.8 \le Z \le 2)$$

$$= \psi(2) + \psi(0.8) = 0.4772 + 0.2881 = 0.7653, \qquad \text{[By } N(0, 1) \text{ tables]}$$

$$p_2 = P(|X - 30| > 5) = P(|Z| > 1) = 2P(Z > 1) = 2(0.5 - 0.3413) = 0.3174$$

$$p_3 = P(X \ge 42) = P(30 + 5Z \ge 42) = P(Z \ge 2.4) = 0.5000 - 0.4918 = 0.0082$$

$$p_4 = P(X \le 28) = P\{Z \le -0.4\} = 0.5 - 0.1554 = 0.3446. \qquad (\textit{Positive Reflection})$$

4*. Let $Z = (X - \mu)/\sigma$, i.e. $X = \mu + \sigma_Z$; $Z \sim N(0, 1)$. Draw a normal curve depiciting regions $A, B, ..., F$.

$$P(A) = P(X > \mu + \sigma) = P(Z > 1) = 0.500 - \psi(1) = 0.50 - 0.3413 = .1587.$$

$$P(B) = P(\mu < X < \mu + \sigma) = P(0 < Z < 1) = \psi(1) = 0.3413.$$

$$P(C) = P(\mu - \sigma < X < \mu) = P(-1 < Z < 0) = \psi\,(1) = 0.3413$$

$$P(D) = P(\mu - 2\sigma < X < \mu - \sigma) = P(-2 < Z < -1) = \psi(2) - \psi(1) = 0.4772 - 0.3413 = 0.1359.$$

$$P(F) = P(X < \mu - 2\sigma) = P(Z < -2) = P(Z > 2) = 0.5 - \psi(2) = 0.0228.$$

Note. $P(A) + P(B) + P(C) + P(D) + P(F) = 1$.

5*. $p = P\{Y \le 3137\} = P\{X^2 + 1 \le 3137\} = P\{X^2 \le 3136\} = P\{-56 \le X \le 56\}$

$$= P\{-10.6 \le Z \le 0.6\} = \psi\,(-10.6) + \psi\,(0.6) = 0.5000 + 0.2258 = 0.7258. \quad \text{[draw Fig.]}$$

6*. Put $Z = (X - \mu)/\sigma = (X - 1)/2$, then $X = 1 + 2Z$. Now

$$p_1 = P\{\tfrac{1}{4} < Z < \tfrac{3}{4} \mid Z > -\tfrac{1}{2}\} = \frac{P(1/4 < Z < 3/4)}{P(Z > -1/2)} = \frac{\psi(3/4) - \psi(1/4)}{0.5 + \psi\,(1/2)}$$

$$= (0.2734 - 0.0987)/(0.5 + 0.1915) = 1747/6915 = 0.2526.$$

$$p_2 = P\{Z < -1/2 \mid |Z| > 1/4\} = P\{|Z| > 1/4, Z < -1/2\}/\{1 - P(|Z| < 1/4)\}$$

$$= P\{Z < -1/2\}/\{1 - P(|Z| < 1/4)\} = [0.5 - \psi\,(1/2)]/[1 - 2\,\psi(1/4)]$$

$$= (0.50 - 0.1915)/(1 - 2 \times 0.0987) = 0.3085/0.8026 = 0.3844.$$

7*. Let $g(x) = P\{x < X < x+\tau\} = \int_x^{x+\tau} f(z)\,dz$; $f(x) = \dfrac{1}{\sigma\sqrt{2\pi}} \exp\left[-\dfrac{(x-\mu)^2}{2\sigma^2}\right]$.

For max. of $g(x)$, $g'(x) = 0$ and $g''(x) < 0$. Now

$g'(x) = f(x+\tau) - f(x)$; $g''(x) = \dfrac{1}{\sigma^2}\{(x-\mu)f(x) - (x+\tau-\mu)f(x+\tau)\}$

$g'(x) = 0 \Rightarrow (x+\tau-\mu)^2 = (x-\mu)^2 \Rightarrow x = \mu - \frac{1}{2}\tau.\ g''(\mu - \frac{1}{2}\tau) = -[\tau e^{-\tau^2/8\sigma^2} / \sigma^2\sqrt{2\pi}] < 0.$

Thus, for $x = \mu - (\tau/2)$. the probability $P(x \le X \le x+\tau)$ is maximum.

Sec. 16-31. Page 494

1*. If σ_A and σ_B are the S.D.'s of Universes A and B, then we are given that $\sigma_A = k\sigma_B$. If y_A and y_B are the maximum frequencies of A and B, and if N is the total frequency then [§16.30(4)]

$$y_A = \frac{N}{\sqrt{2\pi}\,\sigma_A},\quad y_B = \frac{N}{\sqrt{2\pi}\,\sigma_B} = k\frac{N}{\sqrt{2\pi}\,\sigma_A} = ky_A \Rightarrow y_A = \frac{y_B}{k}.$$

2*. We are given that $\mu = 50$, P.E. = 13.49. Now

$Q_1 = \mu - \text{P.E.} = 50 - 13.49 = 36.51.\qquad Q_2 = \mu + \text{P.E.} = 50 + 13.49 = 63.49$

$\text{P.E.} = (2/3)\,\sigma \Rightarrow \sigma = (3/2)\,(13.49) = \frac{1}{2}(40.47) = 20.235.$

$$\text{M.D.} = (4/5)\sigma = (4/5)\,(20.235) = 16.188.$$

$$\text{Mode} = \text{Median} = \text{Mean} \Rightarrow \text{Mode} = 50 = Q_2 \text{ (median)}.$$

Now Cum. frequency = $1250 = \frac{1}{4}(5000) = \frac{1}{4}$ total frequency .

Hence, the variate value must correspond to the lower quartile Q_1. That is $Q_1 = 36.51$.

Sec. 16-41. Page 496

1*. We put $Y = kX^2$; then

$$M(t:Y) = \mathrm{E}(e^{tY}) = \mathrm{E}(e^{tkX^2}) = \int_{-\infty}^{\infty} e^{ktx^2}\frac{e^{-x^2/2}}{\sqrt{2\pi}}\,dx = \frac{1}{\sqrt{1-2kt}}\int_{-\infty}^{\infty}\frac{e^{-u^2}}{\sqrt{2\pi}}\,du, \qquad [\sqrt{1-2kt}\;x = u]$$

$$= (1-2kt)^{-1/2}.$$

Thus $Y \sim \text{gam}\left(\dfrac{1}{2}, \dfrac{1}{2k}\right)$.

Recall : $Y \sim$ gam (a, λ), then $cY \sim \text{gam}(a, \lambda/c)$. Now Y is $\text{gam}\left(\dfrac{1}{2}, \dfrac{1}{2k}\right)$, hence $Z = 2Y$ is $\text{gam}\left(\dfrac{1}{2}, \dfrac{1}{4k}\right)$.

2*. For $N(\mu, \sigma^2)$ we have $M(t : X) = \exp\,(\mu t + \frac{1}{2}\sigma^2 t^2)$. Hence comparing $\mu = 3$, $\sigma = 4$, so given variate X is $N(3, 4^2)$. Now, letting $Z = (X-3)/4$, i.e. $X = 3 + 4Z$.

$$p = P(-1 < X < 9) = P(-1 < Z < 3/2)$$

$$= \psi(3/2) + \psi(1) = 0.4332 + 0.3413 = 0.7745. \qquad \text{[Draw Figure]}$$

3*. Here $\varphi(x) = (2\pi)^{-1}\exp(-x^2/2), \qquad f(x) = (2\pi)^{-1/2}\exp[-(x-\mu)^2/2].$

$\therefore \qquad f(x)/\varphi(x) = \exp(-\mu^2/2)\,.\,e^{\mu x}$. Now

$$\mathsf{E}(Y) = \int_{-\infty}^{\infty} \frac{1-\Phi(x)}{\varphi(x)} f(x)\,dx = e^{-\mu^2/2}\int_{-\infty}^{\infty} [1-\Phi(x)]e^{\mu x}\,dx.$$

Integrating by parts now yields

$$e^{\mu^2/2}\mathsf{E}(Y) = \left\{[1-\Phi(x)]\frac{e^{\mu x}}{\mu}\right\}_{-\infty}^{\infty} + \frac{1}{\mu}\int_{-\infty}^{\infty} e^{\mu x}\varphi(x)\,dx$$

$$= 0 + \mu^{-1}\,[\text{m.g.f. of } N(0, 1) \text{ with } t = \mu] = \mu^{-1}e^{\mu^2/2}.$$

Thus, $\mathsf{E}(Y) = \mu^{-1}$.

4*. $M(t : W) = \mathsf{E}(e^{tW}) = \mathsf{E}_z\{\mathsf{E}(e^{tW} \mid Z = z\}$ [By Double-E Rule}

$$= \mathsf{E}[\mathsf{E}(e^{tX/\sqrt{1+z^2}})\,.\,\mathsf{E}(e^{tYz/\sqrt{1+z^2}}) \mid Z = z] = \mathsf{E}[e^{\frac{1}{2}t^2/(1+Z^2)}\,.\,e^{\frac{1}{2}t^2Z^2/(1+Z^2)}] = \mathsf{E}(e^{t^2/2}) = e^{t^2/2}$$

By Continuity Theorem for m.g.f. we have $W \sim N(0, 1)$.

5*. Let $Y = a - X$. Then, using $\mathsf{E}(X) = \mathsf{E}(Y)$, we get $\mu = a - \mu \;\Rightarrow\; a = 2\mu$. Also

$$\text{Var}(Y) = \text{Var}(a - X) = \text{Var}(X) = \sigma^2.$$

Further, $\qquad \text{Cov}(X, Y) = \text{Cov}(X, a - X) = -\text{Var}(X) < 0.$ So $\rho_{X,Y} < 0$.

Thus, $Y = 2\mu - X$ is the r.v. asked for.

6*. Here $\mathsf{E}(X) = \mu = 0$, $[r = 1]$, hence the given conditions amount to

$$\mu_{2r} = 2r!/2^r\,(r!) \text{ and } \mu_{2r-1} = 0.$$

$$M(t) = \sum_{r=0}^{\infty}\frac{t^r}{r!}\mu_r = \sum_{r=0}^{\infty}\frac{t^{2r}}{(2r)!}\mu_{2r} = \sum_{r=0}^{\infty}\frac{(t^2/2)^r}{r!} = e^{t^2/2}.$$

This is the m.g.f. of $N(0, 1)$, hence variate X has p.d.f. $f(x) = (\sqrt{2\pi})^{-1}e^{-x^2/2}, -\infty < x < \infty$.

7*. For $N(0, \sigma^2)$, $f(x) = (\sigma\sqrt{2\pi})^{-1}e^{-x^2/2\sigma^2}$, hence by comparison, we conclude that $X \sim N(0, 1/2h^2)$ and Y is $N(0, 1/2k^2)$ and so $M(t : X) = e^{t^2/4h^2}$. etc. Now

$$M(t : U) = M(t : X + Y) = M(t : X)\,M(t : Y) = e^{t^2/4h^2}\,.\,e^{t^2/4k^2} = e^{(h^{-2}+k^{-2})t^2/4} = e^{t^2/4l^2} \quad \text{[by indep.]}$$

This shows that $U \sim N(0, 1/2l^2)$ and $\varphi(u) = (l/\sqrt{\pi})e^{-l^2u^2}$.

Note. Result follows trivially from Reproductive Property.

8*.
$$M(it : (X-a)^2) = \int_{-\infty}^{\infty}\frac{e^{-x^2/2}}{\sqrt{2\pi}}e^{it(x-a)^2}\,dx, \qquad \text{(Def.)}$$

$$= \frac{e^{\theta a^2}}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\exp\left\{-\frac{1}{2}(1-2\theta)\left(x^2 + \frac{4a\theta}{1-2\theta}x\right)\right\}dx \quad (\theta = it)$$

$$= \exp\left(\theta a^2 + \frac{2a^2\theta^2}{1-2\theta}\right)\int \exp\left\{-\left(\frac{1-2\theta}{2}\right)\left(x + \frac{2a\theta}{1-2\theta}\right)^2\right\}\frac{dx}{\sqrt{2\pi}}$$

$$M(it) = \exp\left(\frac{\theta a^2}{1-2\theta}\right)\int_{-\infty}^{\infty}\frac{e^{-\frac{1}{2}z^2}}{\sqrt{2\pi}}\frac{dz}{\sqrt{1-2\theta}} \qquad \left\{z = (1-2\theta)^{1/2}\left(x + \frac{2a\theta}{1-2\theta}\right)\right\}$$

$$= (1-2\theta)^{-1/2}.\exp\left[\theta a^2/(1-2\theta)\right].$$

$$\therefore \quad K(t) = \ln M(it) = [\theta a^2/(1-2\theta)] - \tfrac{1}{2}\ln(1-2\theta) = a^2\sum_{r=0}^{\infty}(2\theta)^r + \frac{1}{2}\sum_{r=0}^{\infty}\frac{(2\theta)^r}{r}$$

$$\therefore \quad k_r = \text{Coeff. of } (\theta^r/r!) = 2^{r-1}a^2 r! + 2^{r-1}(r-1)! = 2^{r-1}(r-1)!(1+ra^2).$$

Sec. 16-51. Page 498

1*. From $E(X-10)' = 40$ follows $E(X) = 10 + 40 = 50$, i.e. $\mu = 50$.
Thus $E(X-50)^4 = 48 \Rightarrow \mu_4 = 48$. But $\mu_4 = 3\sigma^4$, hence $3\sigma^4 = 48 \Rightarrow \sigma = 2$. Thus $X \sim N(50, 4)$.

2*. $$E(Y) = (\tfrac{1}{2}\sigma^{-2})E(X-\mu)^2 = (\tfrac{1}{2}\sigma^{-2})\sigma^2 = \tfrac{1}{2}.$$

$$E(Y^2) = E\{(X-\mu)^4/4\sigma^4\} = (1/4\sigma^4)E(X-\mu)^4 = \mu_4/4\sigma^4 = 3/4$$

since for $N(\mu, \sigma^2), \mu_4 = 3\sigma^4$. Thus $\text{Var}(Y) = E(Y^2) - E^2(Y) = (3/4) - (1/4) = 1/2$.

Sec. 16-56. Page 501

1*. Let T_A and T_B be the total travel time for the cars A and B.
Then with obvious meanings of the symbols, $T_A = T_{41} + T_{12}$, so

$$E(T_A) = E(T_{41}) + E(T_{12}) = 3 + 5 = 8 \text{ hrs.} \qquad [\text{c.v.} = \sigma/\mu = 0.20 \Rightarrow \sigma = 0.2\,\mu]$$

$$\text{Var}(T_A) = \sigma_{41}^2 + \sigma_{12}^2 = (0.2\times 3)^2 + (0.2\times 5)^2 = 0.36 + 1 = 1.36.$$

(i) $$P\{T_A < 9\} = \Phi[(9-8)/\sqrt{1.36}] = \Phi(1/1.1662) = \Phi(0.8575) = 0.805.$$

(ii) $$T_B = T_{43} + T_{32},\ E(T_B) = E(T_{43}) + E(T_{32}) = 4 + 2 = 6 \text{ hrs.}$$

$$\text{Var}(T_B) = \sigma_{43}^2 + \sigma_{32}^2 = (0.2\times 4)^2 + (0.2\times 2)^2 = 0.64 + 0.16 = 0.80.$$

Let $Z = T_A - T_B$, then $Z \sim N(8-6, 1.36+0.80)$ i.e. $Z \sim N(2, 2.16)$.
$P\{Z < 0\} = \Phi[(0-2)/\sqrt{2.16}] = \Phi(-2/1.4697) = \Phi(-1.361) = 0.087$.

2*. The mean and variance of the new (composite) distribution are $\mu = \mu_x + \mu_y = 23$; $\sigma^2 = \sigma_x^2 + \sigma_y^2 = 5.76 + 9 = 14.76$, $\sigma = \sqrt{14.76} = 3.84$. Now 1/1000 limits (i.e. 1 in 1000) means that the area cut from each tail is 0.001. The critical point corresponding to this tail area is, assuming normality, $k = 3.09$. The limits of our interest are $\mu \pm k\sigma$, which are $23 \pm 3.09 \times 3.84 = 23 \pm 11.87$. Thus, the limits are 34.87 and 11.13.

3*. Denote the marks in the given subjects by X, Y, Z then the total marks are given by $T = X + Y + Z$. Now,

$$E(T) = E(X + Y + Z) = 150, \text{Var}(T) = \text{Var}(X) + \text{Var}(Y) + \text{Var}(Z) = 625.$$

Thus T is $N(150, 25^2)$ by additive property. Now using $N(0, 1)$-Tables, we get

(a) $P(T \geq 180) = P\{(T - 150)/25 \geq 1.2\} = 0.1942$

(b) $P(T \leq 90) = P\{(T - 150/25 \leq -2.40\} = 0.0224.$

4*. Let $U = 3X + Y$, then $E(U) = 0$, $\text{Var}(U) = 9\sigma_x^2 + \sigma_y^2 = 90.$

Let D be the region bounded by the parallel st. lines $3X + Y = 10$ and $3X + Y = 5$, i.e. $5 \leq U \leq 10$. Now $Z = (U - 0)/\sqrt{90} \sim N(0, 1)$. $[U = \sqrt{90}\, Z.]$

$$\therefore \quad p = P\{(X,Y) \in D\} = P\{5 < U < 10\} = P\{(5/\sqrt{90}) < Z < 10/\sqrt{90}\}$$

$$= P\{0.53 < Z < 1.05\} = \psi(1.05) - \psi(0.53) = 0.3531 - 0.2019 = 0.1512. \quad \text{[Draw Fig.]}$$

5*. $p = P\{1.202 < X/Y < 8318 \times 10^4\} = P\{\ln 1.202 < (\ln X - \ln Y) < 4 + \ln 8318\}$
$= P\{0.08 < U < 7.92\}$ ...(1)

where $U = (\ln X - \ln Y) \sim N(7 - 3, 3 + 1) = N(4, 4)$ by *closure* property of $N(\mu, \sigma^2)$.

Thus letting $Z = (U - 4)/2$ i.e. $U = 4 + 2Z$, we get from (1)

$$p = P\{0.08 < 4 + 2Z < 7.92\} = P\{-1.96 < Z < 1.96\} = 0.95. \quad [Z \sim N(0, 1)].$$

6*. $\mu = E(Y) = \sum c_i E(X_i) = \sum c_i \mu_i$; $\text{Var}(Y) = \sum c_i^2 \sigma_i^2 = \sigma^2$ (say). Thus, by hypothesis $9\sigma^2 = \mu^2$, i.e. $\sigma = \mu/3$.

By Closure Property, $Y \sim N(\mu, \sigma^2)$. Put $Z = (Y - \mu)/\sigma$, i.e. $Y = \mu + (\mu/3)Z$

$P\{0 \leq Y \leq 2\mu\} = P(0 \leq \mu + (\mu/3) Z \leq 2\mu) = P\{-3 \leq Z \leq 3\} = 0.9973.$ [Area under 3-σ limits].

Sec. 16-73. Page 512

1*. Here $\varphi(x) = (\sqrt{2\pi})^{-1} \exp(-x^2/2)$. Put $\lambda = 1/[\Phi(b) - \Phi(a)]$. Now

$$E(Y) = \int_a^b y g(y)\, dy = \lambda \int_a^b y\varphi(y)\, dy = \frac{\lambda}{\sqrt{2\pi}} \int_a^b y e^{-y^2/2} dy = \frac{\lambda(e^{-a^2/2} - e^{-b^2/2})}{\sqrt{2\pi}} = \frac{\varphi(a) - \Phi(b)}{\varphi(b) - \Phi(b)}$$

Converse. Let $g(y) = f(y)/[F(b) - F(a)] = k f(y)$, where $k = [F(b) - F(a)]^{-1}$. $a < y < b$ be the p.d.f. of a truncated variate Y, alongwith $E(Y) = K[f(a) - f(b)]$.

$$\therefore \quad \frac{f(a) - f(b)}{F(b) - F(a)} = \int_a^b y\, g(y)\, dy = \int_a^b \frac{y f(y)\, dy}{F(b) - F(a)} \Rightarrow f(a) - f(b) = \int_a^b y f(y)\, dy.$$

Differentiating partially w.r.t. 'b' under integral sign : $-f'(b) = bf(b) \Rightarrow f'(b)/f(b) = -b$.

Integrating : $\ln f(b) = -\frac{1}{2} b^2 + \text{constant} \Rightarrow f(b) = Ke^{-b^2/2}$

where K is a constant independent of b (may be function of a).

Since 'f' is a p.d.f. $\int_{-\infty}^{\infty} f(b)\, db = 1 \Rightarrow K = 1/\sqrt{2\pi}$. Hence $f(b) = (\sqrt{2\pi})^{-1} e^{-b^2/2}$, $-\infty < b < \infty$.

Had we differentiated (1) partially w.r.t. 'a', we would have got

$$f(a) = (\sqrt{2\pi})^{-1} e^{-a^2/2}, \quad -\infty < a < \infty.$$

It follows that 'f' is the density function of $N(0, 1)$.

2*. (i) Let $X \sim N(0, 1)$ and $Y \sim N(0, 1)$ be independent variates. By Double-E Rule,

$$\mathsf{E}(e^{itXY}) = \mathsf{E}[\mathsf{E}_Y(e^{itXY} \mid X = x)] = \mathsf{E}_X[\mathsf{E}_Y(e^{itxY})] = \mathsf{E}(e^{t^2X^2/2})$$

$$= \int_{-\infty}^{\infty} e^{t^2x^2/2} \cdot \frac{e^{-x^2/2}}{\sqrt{2\pi}}\, dx = \int_{-\infty}^{\infty} \frac{e^{-(1-t^2)x^2/2}}{\sqrt{2\pi}}\, dx = \int_{-\infty}^{\infty} \frac{e^{-z^2/2} dz}{\sqrt{2\pi} \cdot \sqrt{1-t^2}} \quad (z = x\sqrt{1-t^2})$$

$$= (1-t^2)^{-1/2} \text{ [since area under } N(0, 1)\text{-curve is unity]}$$

As $(1-t^2)^{-1/2}$ can never be expressed as $e^{t^2/2}$, it follows that XY can never be a normal variate.

Note. $\mathsf{E}(e^{itXY})$ can be evaluated without conditioning ; but conditioning speeds up evaluation.

(ii) Let X and Y be independent variates with p.d.fs.

$$f_1(x) = 1/\pi\sqrt{1-x^2}, |x| < \infty; \qquad f_2(y) = (y/\sigma^2)\, e^{-y^2/2\sigma^2}, \; y > 0$$

The p.d.f. of $Z = XY$ is given by the standard formula

$$f_z(z) = \int_{-\infty}^{\infty} \frac{1}{|w|} f_1\left(\frac{z}{w}\right) f_2(w)\, dw$$

Now, $f_2(w) = 0$ for $w < 0$, and $f_1(z/w) = 0$ for $0 < w < |z|$. Hence

$$f_z(z) = \int_{|z|}^{\infty} \frac{we^{-w^2/2\sigma^2}\, dw}{w\sigma^2\pi\sqrt{1-(z^2/w^2)}} = \frac{1}{\pi\sigma^2} \int_{|z|}^{\infty} \frac{we^{-w^2/2\sigma^2}\, dw}{\sqrt{w^2 - z^2}}.$$

Let $w^2 - z^2 = 2\sigma^2 t$, then $w\, dw = \sigma^2\, dt$ and we obtain

$$f_z(z) = \frac{e^{-z^2/2\sigma^2}}{\pi\sigma\sqrt{2}} \int_0^{\infty} t^{1/2-1} e^{-t}\, dt = \frac{e^{-z^2/2\sigma^2}}{\sigma\sqrt{2\pi}} \qquad \left[\because \Gamma\left(\tfrac{1}{2}\right) = \sqrt{\pi}\right]$$

It follows that $Z = XY \sim N(0, \sigma^2)$.

3*. For each n, let $g_n(x)$ denote the p.d.f. of X_n and $F'(x) = f(x)$. If $h(x)$ is the p.d.f. of $N(0, 1)$, differentiating the given relation, we get

$$g_n(x) = [(n-1)/n]\, h(x) + n^{-1} f_n(x), \quad \forall\, n$$

Consider the Ch. Function of X_n,

$$\phi(t : X_n) = \mathsf{E}(e^{itX_n}) = \int_{-\infty}^{\infty} e^{itx} \cdot g_n(x)\, dx = \left(\frac{n-1}{n}\right) \int_{-\infty}^{\infty} e^{itx} h(x)\, dx + \frac{1}{n} \int_{-\infty}^{\infty} e^{itx} f_n(x)\, dx$$

$$= [(n-1)/n]\, e^{-t^2/2} + (1/n)\, \phi(t : Y_n) \qquad \text{...(1)}$$

Since $|\phi(t:Y_n)| \le 1$, $\lim (1/n)\,\phi(t:Y_n) \to 0$, as $n \to \infty$.

$$\therefore \quad \lim \phi(t:X_n) = \lim (1-n^{-1})\,e^{-t^2/2} + \lim (1/n)\,\phi(t:Y_n) = e^{-t^2/2} = \text{Ch. Fun. of } N(0,1)$$

By Continuity Theorem, limiting distribution of X_n is $N(0, 1)$.

4*. $F_Y(y) = P(Y \le y) = P(X^2 \ge 1/y) = 1 - P(X^2 \le y^{-1}) = 1 - P(-y^{-1/2} \le X \le y^{-\frac{1}{2}})$

$$= 1 - 2\,P(0 \le X \le y^{-\frac{1}{2}}) = 1 - 2\int_0^{1/\sqrt{y}} f(x)\,dx \,. \left[f(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-(x^2/2\sigma^2)} \right]$$

By DUIS we get,

$$f_Y(y) = -2\frac{\partial}{\partial y}\int_0^{1/\sqrt{y}} f(x)\,dx = \frac{1}{y^{3/2}}\, f_X\!\left(\frac{1}{\sqrt{y}}\right) = (\sigma\sqrt{2\pi})^{-1}\, y^{-3/2}\, e^{-1/2\sigma^2 y}, \quad 0 \le y < \infty.$$

$$E(e^{-tY}) = \int_0^\infty e^{-ty} \,.\, f_Y(y)\,dy = \int_0^\infty e^{-t/x^2} \,.\, \frac{2e^{-x^2/2\sigma^2}}{\sigma\sqrt{2\pi}}\,dx \qquad \left[y = \frac{1}{x^2} \right]$$

$$= \frac{2}{\sigma\sqrt{2\pi}} \int_0^\infty e^{-(x^2/2\sigma^2)-(t/x^2)}\,dx, \qquad \left[\int_0^\infty e^{-a^2x^2-(b^2/x^2)}\,dx = \frac{\sqrt{\pi}\,e^{-2ab}}{2a} \right] \qquad \ldots(1)$$

$$= \frac{2}{\sigma\sqrt{2\pi}} \cdot \frac{\sqrt{\pi}}{(2/\sqrt{2\sigma^2})}\, e^{-2\sqrt{t}/\sqrt{2\sigma^2}} = e^{-\sqrt{2t}/\sigma}$$

$$\therefore\ \phi_Y(t) = E(e^{itY}) = e^{-\sqrt{-2it}/\sigma}$$

Note. $$E(e^{-tY}) = E(e^{-t/X^2}) = \int_{-\infty}^{\infty} e^{-t/x^2}\, \frac{e^{-x^2/2\sigma^2}}{\sigma\sqrt{2\pi}}\,dx = e^{-\sqrt{2t}/\sigma}.$$

This is identical with (1). We have avoided the use of contour integration.

5*. Let $Z = X_1X_2X_3 \big/ \sqrt{X_1^2X_2^2 + X_2^2\,X_3^2 + X_3^2X_1^2}$ then $(1/Z^2) = \Sigma(1/X_i^2)$.

$$\phi(t:1/Z^2) = \phi(t:1/X_1^2)\,\phi(t:1/X_2^2)\,\phi(t:1/X_3^2)$$

$$= e^{-\sqrt{-2it}/\sigma_1}\, e^{-\sqrt{-2it}/\sigma_2}\, e^{-\sqrt{-2it}/\sigma_3} = e^{-\sqrt{-2it}/\sigma} \qquad \text{[by 4*]}$$

Thus Ex. 4* shows that $W = 1/Z^2$ has p.d.f. $f(w) = (\sigma\sqrt{2\pi})^{-1}\, w^{-3/2}\, e^{-1/2\sigma^2 w}, w \ge 0$.

The p.d.f. $f(z)$ of Z is given by

$$g(z) = f(w)\left|\frac{dw}{dz}\right| = \frac{1}{\sigma\sqrt{2\pi}} \,.\, e^{-z^2/2\sigma^2}, \ -\infty < z < \infty.$$

Extension. If $X_k \sim N(0, \sigma_k^2)$, $k = 1, 2, \ldots, n$ are independent, then

$$Z = X_1\,X_2 \ldots X_n / \sqrt{T}, \ \text{ where } T = X_1^2X_2^2 + X_2^2X_3^2 + \ldots + \ldots + X_{n-1}^2\,X_n^2.$$

is $N(0, \sigma^2)$, where $1/\sigma = \Sigma(1/\sigma_k)$.

6*. Let $Z = (X - \mu)/\sigma$, i.e. $X = \mu + \sigma Z$. Now

$$M(t:V) = E(e^{tV}) = E(e^{tX^2/\sigma^2}) = \int_{-\infty}^{\infty} e^{tx^2/\sigma^2} \cdot \frac{e^{-(x-\mu)^2/2\sigma^2}}{\sqrt{2\pi}\,\sigma}\,dx = \int_{-\infty}^{\infty} e^{-z^2/2} \cdot e^{t(\mu+\sigma z)^2/\sigma^2}\,dz \qquad \ldots(1)$$

$$\text{Exponent} = \frac{\mu^2 t}{\sigma^2} - \left(\frac{1-2t}{2}\right)\left[z^2 - \frac{4\mu t z}{(1-2t)\sigma} + \frac{4\mu^2 t^2}{(1-2t)^2\sigma^2}\right] + \frac{2\mu^2 t^2}{(1-2t)\sigma^2}$$

$$= \frac{\mu^2 t}{(1-2t)\sigma^2} - \left(\frac{1-2t}{2}\right)\left[z - \frac{2\mu t}{(1-2t)\sigma}\right]^2 = \frac{\mu^2 t}{(1-2t)\sigma^2} - \frac{1}{2}u^2,$$

where $u = \sqrt{1-2t}\left[z - \dfrac{2\mu t}{(1-2t)\sigma}\right]$

Substituting into (1) this exponent, using $dz = du/\sqrt{1-2t}$ we get

$$M(t:V) = \frac{\exp[\mu^2 t/(1-2t)\sigma^2]}{\sqrt{1-2t}}\left(\int_{-\infty}^{\infty}\frac{e^{-u^2/2}}{\sqrt{2\pi}}\right) = \frac{\exp[2\lambda t/(1-2t)]}{\sqrt{1-2t}}, \qquad \left[\lambda = \frac{\mu^2}{2\sigma^2}\right] \qquad \ldots(2)$$

$$= (1-2t)^{-1/2} \cdot e^{-\lambda} \cdot e^{\lambda(1-2t)^{-1}} \qquad [2t/(1-2t) = (1-2t)^{-1} - 1]$$

$$= (1-2t)^{-1/2} \cdot e^{-\lambda} \sum_{r=0}^{\infty} \frac{\lambda^r (1-2t)^{-r}}{r!} = \sum_{r=0}^{\infty} \frac{e^{-\lambda}\lambda^r}{r!}(1-2t)^{-(1+2r)/2} \qquad \ldots(3)$$

$$\therefore \qquad \phi(t:V) = \sum_{r=0}^{\infty} \frac{e^{-\lambda}\lambda^r}{r!}(1-2it)^{-(1+2r)/2} = \sum_{r=0}^{\infty} p_r (1-2it)^{-(1+2r)/2}. \qquad \left[p_r = \frac{e^{-\lambda}\lambda^r}{r!}\right] \qquad \ldots(4)$$

To find density of V we use Inversion formula to get

$$f(V) = \frac{1}{2\pi}\int_0^{\infty} \phi(t)\, e^{-itv}\,dt = \sum_{r=0}^{\infty} p_r \left\{\frac{1}{2\pi}\int_0^{\infty} e^{-itv}(1-2it)^{(1+2r)/2}\,dt\right\} \qquad \ldots(5)$$

$$\text{Recall : For } Y \sim \chi^2_{(1+2r)},\ g(y) = \frac{e^{-y/2} \cdot y^{(1+2r)/2-1}}{\Gamma[(1+2r)/2]\,2^{(1+2r)/2}},\ \phi_Y(t) = (1-2it)^{-(1+2r)/2} \qquad \ldots(6)$$

Using (6) we instantly obtain

$$f(v) = \sum_{r=0}^{\infty} p_r\, g(y) = \sum_{r=0}^{\infty} p_r \frac{e^{-v/2}(v)^{r-(1/2)}}{\Gamma(r+1/2)^{r+(1/2)}} = \frac{e^{-\lambda} v^{-1/2}\, e^{-v/2}}{\sqrt{2\pi}} + \sum_{r=1}^{\infty} p_r \frac{e^{-v/2}\, v^{r-1}}{2^{r+1/2}\,\Gamma(r+1/2)}. \qquad \ldots(7)$$

7*. Since X is $N(\mu, \sigma^2)$, $f(x) = (\sigma\sqrt{2\pi})^{-1}\exp\{-(x-\mu)^2/2\sigma^2\}$. Let $z = (x-\mu)/\sigma$ i.e. $x = \mu + \sigma z$. The total probability of the event A, i.e. $P(A)$ is given by the Partitioning Theorem :

$$P(A) = \int_{-\infty}^{\infty} f(x)\,P(A \mid x)\,dx = \int_0^{\infty} \frac{e^{-(x-\mu)^2/2\sigma^2}}{\sigma\sqrt{2\pi}}(1 - e^{-kx})\,dx,$$

$$= \int_{-\mu/\sigma}^{\infty} \frac{e^{-z^2/2}}{\sqrt{2\pi}}\,dz - \int_{-\mu/\sigma}^{\infty} \frac{\exp[-\frac{1}{2}(z+k\sigma)^2 - \mu k + \frac{1}{2}k^2\sigma^2]}{\sqrt{2\pi}}\,dz,$$

$$= \frac{1}{2} + \psi\left(\frac{\mu}{\sigma}\right) - \exp\left(-\mu k + \frac{1}{2}\sigma^2 k^2\right) \int_{k\sigma - (\mu/\sigma)}^{\infty} \frac{e^{-u^2/2}}{\sqrt{2\pi}}\, du, \quad (u = z + k\sigma)$$

$$= \frac{1}{2} + \psi\left(\frac{\mu}{\sigma}\right) - e^{-\mu k + \frac{1}{2}\sigma^2 k^2}\left[\frac{1}{2} + \psi\left(\frac{\mu - k\sigma^2}{\sigma}\right)\right]. \qquad \left[\psi(k) = \int_0^k \frac{e^{-\frac{1}{2}z^2}}{\sqrt{2\pi}}\, dz.\right]$$

8*. Since $X \sim N(0, \sigma^2)$, $E(X) = E(X^3) = 0$, $E(X^2) = \sigma^2$, $(X^4) = 3\sigma^4$

$Z = X + Y$, $E(Z) = E(X) + E(Y) = 0$

$$E(Z^2) = E(X^2 + Y^2 + 2XY) = \sigma^2 + E(Y^2) + 2E(X)E(Y) = \sigma^2 + \sigma_Y^2$$

$$3[E(Z^2)]^2 = 3(\sigma^4 + \sigma_Y^4 + 2\sigma^2\sigma_Y^2) = 3\sigma^4 + E(Y^4) + 6\sigma^2 E(Y^2). \quad [E(Y^4) = 3\sigma_Y^4] \quad \ldots(1)$$

$$E(Z^4) = E(X+Y)^4 = E(X^4 + 4X^3Y + 6X^2Y^2 + 4XY^3 + Y^4)$$

$$= E(X^4) + 4E(X^3)E(Y) + 6E(X^2)E(Y^2) + 4E(X)E(Y^3) + E(Y^4)$$

$$= 3\sigma^4 + 0 + 6\sigma^2 E(Y^2) + 0 + E(Y^4) = 3[E(Z^2)]^2, \quad \text{[by (1)]} \quad \ldots(2)$$

9*. Here, $f(x) = (\sqrt{2\pi})^{-1} e^{-\frac{1}{2}x^2}$, $f'(x) = -xf(x)$, $\Phi'(x) = f(x)$ $\ldots(1)$

$$\therefore \quad \frac{d}{dx}\left[\left(\frac{1}{x} - \frac{1}{x^3}\right)f(x)\right] = \left(\frac{1}{x} - \frac{1}{x^3}\right)f'(x) + \left(\frac{3}{x^4} - \frac{1}{x^2}\right)f(x) = \left(\frac{3}{x^4} - 1\right)f(x), \quad \text{[by (1)]}$$

$$\frac{d}{dx}\left[\frac{f(x)}{x}\right] = \frac{f'(x)}{x} - \frac{f(x)}{x^2} = -\left(1 + \frac{1}{x^2}\right)f(x), \quad \text{[by (1)]}$$

Integrate (1) and (2) in the range (x, ∞), use $f(\infty) = 0$, to get

$$\left(\frac{1}{x} - \frac{1}{x^3}\right)f(x) = \int_x^{\infty}\left(1 - \frac{3}{y^4}\right)f(y)\,dy; \quad \frac{f(x)}{x} = \int_x^{\infty}\left(1 + \frac{1}{y^2}\right)f(y)\,dy$$

Since $1 - (3/y^4) < 1$, $(1 + y^{-2}) > 1$, these relations yield

$$\left(\frac{1}{x} - \frac{1}{x^3}\right)f(x) \le \int_x^{\infty} f(y)\,dy = 1 - \Phi(x); \quad \frac{f(x)}{x} \ge \int_x^{\infty} f(y)\,dy = 1 - \Phi(x).$$

Combining these relations we get the stated result (a). For large x, these bounds are clearly very tight.

(b) We multiply the inequality (a) by $x/f(x) > 0$, to obtain

$$\left(1 - \frac{1}{x^2}\right) \le \frac{x[1 - \Phi(x)]}{f(x)} \le 1.$$

Let $x \to \infty$, then $(1/x^2) \to 0$, and by Squeeze-Principle, this gives

$$\lim_{x \to \infty} \frac{x[1 - \Phi(x)]}{f(x)} = 1.$$

Remark. The result (b) provides : $1 - \Phi(x) \sim x^{-1} f(x)$.

Re-statement. $x^{-1}(1 - x^{-2}) \le r(x) \le x^{-1}$; $\lim x.r(x) = 1$

[**Note.** $a(x) \sim b(x)$ for large x means $\lim a(x)/b(x) = 1$, as $x \to \infty$].

10*. We follow m.g.f. technique. The m.g.f. of $(X_1, X_2, ..., X_n)$ is

$$M(t_1, t_2, ..., t_n) = \mathsf{E}(e^{t_1X_1 + t_2X_2 + ... + t_nX_n})$$

$$= \mathsf{E}\{\exp(t_1Y_1 + t_2(Y_1 + Y_2) + t_3(Y_1 + Y_2 + Y_3) + ... + t_n(Y_1 + ... + Y_n))\}$$

$$= \mathsf{E}\{\exp[(t_1 + ... + t_n)Y_1 + (t_1 + ... + t_n)Y_2 + (t_3 + ... + t_n)Y_3 + ... + t_n Y_n]\}$$

$$= \exp\tfrac{1}{2}[(t_1 + ... + t_n)^2 + (t_2 + ... + t_n)^2 + ... + (t_{n-1} + t_n)^2 + t_n^2], \text{ [using m.g.f. of } N(0, 1)].$$

Now put $t_1 = t_2 = ... = t_n = t/n^{3/2}$; then

$$M(t : S_n/n^{3/2}) = \exp\{\tfrac{1}{2}(t^2/n^3)[n^2 + (n-1)^2 + ... + 2^2 + 1^2]\} = \exp[(n+1)(2n+1)/12n^2]t^2.$$

As $n \to \infty$, $M(t : S_n/n^{3/2}) \to \exp[\frac{1}{2}(t^2/3)]$; so $S_n/n^{3/2} \to N(0, \sigma^2)$, where $\sigma^2 = 1/3$.

11*. From $X_0 = Y_0$, $X_1 + aY_0 + Y_1$, $X_2 = a^2Y_0 + aY_1 + Y_2$, etc., we find that X_n is a linear combination of $Y_0, Y_1, ..., Y_n$ which are indep. Gaussian variates. Hence, X_n is a Gaussian variate. And $\mathsf{E}(Y_j) = 0$, $\forall_j$ provides $\mathsf{E}(X_n) = 0$, $\forall n$. Further,

$$\mathsf{E}(X_n^2) = \mathsf{E}(a_{n-1} + Y_n)^2 = a^2\,\mathsf{E}(X_{n-1})^2 + 2a\,\mathsf{E}(X_{n-1}Y_n) + \mathsf{E}(Y_n^2) \qquad ...(1)$$

Expressing X_{n-1} in terms of Y_j noting $\mathsf{E}(Y_i Y_j) = 0$ $(i < j)$, we find that $\mathsf{E}(X_{n-1}Y_n) = 0$.

$$\therefore \quad \mathsf{E}(X_n^2) = a^2\,\mathsf{E}(X_{n-1})^2 + 1; \text{ hence } \lim \mathsf{E}(X_n^2) = a^2 \lim \mathsf{E}(X_{n-1}^2) + 1 \text{ (as } n \to \infty).$$

This gives : $\lim\limits_{n\to\infty} \mathsf{E}(X_n^2) = \lim\limits_{n\to\infty} \sigma_n^2 = 1/(1-a^2) = \sigma^2$.

Now $X_n \sim N(0, \sigma_n^2)$ so $M(t : X_n) = \exp[\frac{1}{2}\sigma_n^2 t^2]$

$$\lim_{n\to\infty} M(t : X_n) = \lim_{n\to\infty} \exp[\tfrac{1}{2}\sigma_n^2 t^2] = \exp[\tfrac{1}{2}\sigma^2 t^2]$$

Thus, $X_n \xrightarrow{d} N(0, \sigma^2)$ where $\sigma^2 = 1/(1 - a^2)$.

12*. Write $\lambda = [2\pi J_0(k)]^{-1}$ and set $x - \beta = \theta$, and use exponential expansion to get

$$\mathsf{E}(\cos X) = \lambda \int_0^{2\pi} \cos x\, e^{k\cos(x-\beta)} dx = \lambda \int_{-\beta}^{2\pi-\beta} \cos(\beta + \theta)\, e^{k\cos\theta}\, d\theta$$

$$= \lambda \int_{-\beta}^{2\pi-\beta} \cos(\beta+\theta)\left[\sum_{r=0}^{\infty} \frac{(k\cos\theta)^r}{r!}\right] d\theta = \lambda \sum_{r=1}^{\infty} \frac{k^r}{r!} \int_{-\beta}^{2\pi-\beta} (\cos\beta\cos\theta - \sin\beta\sin\theta)\cos^r\theta\, d\theta$$

Since the integrand is *periodic function* with period 2π, shift by β can be dispensed with.

$$\therefore\ \mathsf{E}(\cos X) = \lambda \sum_{r=0}^{\infty} \frac{k^r}{r!} \int_0^{2\pi} (\cos\beta(\cos\theta)^{r+1} - \sin\beta\cos^r\theta\sin\theta)d\theta = \lambda\cos\beta \sum_{r=0}^{\infty} \frac{k^r}{r!} \int_0^{2\pi} (\cos\theta)^{r+1} d\theta. \quad ...(1)$$

$$\left[\int_0^{2\pi} \sin\theta(\cos\theta)^r d\theta = 0\right]$$

Now $I_{r+1} = \int_0^{2\pi} (\cos\theta)^{r+1}\, d\theta = 2\int_0^{\pi} (\cos\theta)^{r+1}\, d\theta = 2(-1)^{r+1}\int_{-\pi/2}^{\pi/2} (\sin\phi)^{r+1}\, d\phi$, $[\theta = (\pi/2) + \phi]$

If r is even, $I_{r+1} = 0$ (odd integrand), so reject r-even values in (1). When $r\ (= 2n - 1)$ is odd use Duplication Formula.

$$I_{2n} = 4\int_0^{\pi/2} (\sin\phi)^{2n} \cos^0\phi\, d\phi = \frac{4\Gamma(1/2)\,\Gamma(n+1/2)}{2\Gamma(n+1)} = \frac{2\sqrt{\pi}}{n!}\left(\frac{\sqrt{\pi}\,\Gamma(2n)}{2^{2n-1}\,\Gamma(n)}\right)\frac{2n}{2n} = \frac{\pi(2n)!}{2^{2n-1}(n!)^2} \quad \ldots(2)$$

$$\therefore\ \mathrm{E}(\cos X) = \lambda\cos\beta \sum_{r=0}^{\infty} \frac{k^{2r+1}}{(2r+1)!}\int_0^{2\pi} (\cos\theta)^{2r+2}\, d\theta = \lambda\cos\beta \sum_{r=0}^{\infty} \frac{k^{2r+1}}{(2^{2r+1})!}\,\frac{\pi(2r+2)!}{2^{2r+1}[(r+1)!]^2}$$

$[n \to (r+1)$ in (2)$]$

$$= \frac{2\pi\cos\beta}{2\pi J_0(k)} \sum_{r=0}^{\infty} \frac{(k/2)^{2r+1}}{r!(r+1)!} = \cos\beta\,\frac{J_1(k)}{J_0(k)} = \rho\,(\cos\beta). \qquad \left[\rho = \frac{J_1(k)}{J_0(k)}\right]$$

$$\mathrm{E}(\sin X) = \lambda\int_0^{2\pi} \sin x\, e^{k\cos(x-\beta)}\, dx = \lambda\int_{-\beta}^{2\pi-\beta} \sin(\beta+\theta)\left(\sum_{r=0}^{\infty} \frac{(k\cos\theta)^r}{r!}\right) d\theta. \qquad [x - \beta = \theta]$$

$$= \lambda\sum_{r=0}^{\infty} \frac{k^r}{r!}\int_0^{2\pi} (\cos\beta\sin\theta\cos^r\theta + \sin\beta\cos^{r+1}\theta)\, d\theta,$$ [Dropping β from range of periodic functions]

These integrals are evaluated above. Only $\sin\beta$ occurs in place of $\cos\beta$. Hence

$$\mathrm{E}(\sin X) = \rho\sin\beta.$$

[For properties of periodic functions, see p. 417 Real Analysis by the Authors]

Appendix E, Page 520

1*. Let $X \sim N(10, 0.01)$ and set $Z = (X - \mu)/\sigma$ i.e. $X = \mu + \sigma Z = 10 + (Z/10)$. Thus [Draw Figure]

$$p = P(9.9 < X < 10.2) = P(9.9 < 10 + (Z/10) < 10.2) = P(-1 < Z < 2)$$
$$= \psi(2) + \psi(1) = 0.4772 + 0.3413 = 0.8185.$$

2*. Let X be the length of the bed (in cm). Assuming a normal distribution we require to find a value L such that $P(X \le L) = 0.95$.

$$\therefore\quad 0.95 = P(X \le L) = P\{Z \le (L - 172)/8.5\},\ [Z = (X - \mu)/\sigma]$$

Standard values give $(L - 172)/8.5 \pm 1.645$ [draw Fig.] so that

$$L = 172 + (8.5) \times (1.645) = 172 + 13.9825 = 185.98.$$

Hence, if the firm makes beds 186 cm in length, then no more than 5% of men will find themselves too long for the beds.

3*. Let X denote the height of an individual, then by hypothesis $X \sim N(68.22, 10.8)$. Put $Z = (X - 68.22)/\sqrt{10.8}$. If p denotes the probability that an individual is over 6 feet (72″), then

$$p = P(X \ge 72)\ P\{68.22 + Z\sqrt{10.8} \ge 72\} = P(Z \ge 1.15)$$
$$= 0.5000 - \psi(1.15) = 0.5 - 0.3746 = 0.1254 \qquad \text{[Draw Fig.]}$$

It follows that the number of soldiers with height over 6 feet is $n = 1000\ p = 125.4 = 125$.

4*. Let X be $N(\mu, \sigma^2)$, here $\mu = 15$; $\sigma = 3.5$. If n is the total number of observations, then we are given that

$$nP(X > 16.25) = 647 \qquad [Z = (X - \mu)/\sigma] \qquad \ldots(i)$$

$$P\{X > 16.25\} = P\{Z > (16.25 - 15)/3.5\} = P\{Z > 0.357\} \qquad \text{[Draw Fig.]}$$

$$= 0.5000 - \psi(0.36) = 0.5 - 0.1406 = 0.3594.$$

Substituting into (i) we get $n = 647/(0.3594) = 1800.2225 = 1800$.

5*. Let X be $N(\mu, \sigma^2)$ $\mu = 78$, $\sigma = 11$, put $z = (x - \mu)/\sigma$ i.e. $x = 78 + 11z$.

(i) When $x = 90$, $z = (90 - 78)/11 = 12/11 = 1.0909 = 1.09$.

$$\therefore \qquad P(X > 90) = P(Z > 1.09) = 0.50 - \psi\,(1.09) = 0.5000 - 0.3621 + 0.1379.$$

Hence, of the 1000 grades, 137.9 i.e. 138 were above 90%.

(i) The probability of one being among the lowest 10 grades is $10/1000 = 0.01$. This gives

$$z = 2.33 = (x - 78)/11.$$

$\therefore\ x = 78 - 11 \times (2.33) = 78 - 25.63 = 52.37 \Rightarrow$ the highest grade of the lowest 10 is 52%.

(iii) $Q = \text{P.E.} = 2\sigma/3 = 22/3 = 7.3\%$.

(iv) We require 450 grades to exceed μ. The probability of one being among these 450 is $450/1000 = 0.45$. Thus $\psi\,[(x - 78)/11] = 0.45$ which gives

$$(x - 78)/11 = 1.645 \quad \Rightarrow x = 78 + 11 \times (1.645) = 78 + 18.095 = 96.1.$$

This is the highest percentage for the range above the mean. From symmetry, we have 450 grades below the mean and their lowest percentage is $78 - 11 \times (1.645) = 78 - 18.095 = 59.905 = 60\%$. Thus the middle 900 grades lie between 60% and 96.1%.

6*. The number of candidates obtaining class III is 600, which implies that the area to the left of $X = 50$ is 0.6. For $N(\mu, \sigma^2)$. Let $Z = (X - \mu)/\sigma$. The figure reads

$$\psi[(50 - \mu)/\sigma] = 0.10, \ \psi\,[(60 - \mu)/\sigma] = 0.45 \qquad \ldots(1)$$

We are not given variate-values corresponding to these areas. Hence we interpolate linearly to get the required values. Linear interpolation is

$$\frac{A - A_1}{A_2 - A_1} = \frac{z - z_1}{z_2 - z_1} \qquad [A = \text{Area}, z = \text{Variate value}]$$

(i) $$\frac{0.100 - 0.079}{0.118 - 0.079} = \frac{z - 0.2}{3 - 0.2} \Rightarrow z = 0.2 + \frac{0.7}{13} = 0.254, \ \psi(0.254) = 0.10$$

(ii) $$\frac{0.450 - 0.445}{0.455 - 0.445} = \frac{z - 1.6}{1.7 - 1.6} \Rightarrow z = 1.6 + \frac{0.5}{10} = 1.65, \text{ i.e. } \psi(1.65) = 0.45$$

Using these values in (1) we get $(50 - \mu) = \sigma(0.254)$ and $(60 - \mu) = \sigma(1.65)$

Subtraction gives $10 = \sigma(1.396)$ or $\sigma = 10/1.4 = 7.16$.

$$\mu = 60 - (1.65)\,\sigma = 60 - 11.815 = 48.185 = 48.2.$$

The central 50% area under $N(\mu, \sigma^2)$ lies between $\mu \pm 2/3\ \sigma = 48.2 \pm 4.733$ giving the boundaries 53.01 and 43.55. Hence roughly speaking about 500 candidates secure marks between 43 and 53. To be exact.

$$p = P(43 < X < 53) = P[(43 - 48.2)/7.16 < (X - \mu)/\sigma < (53 - 48.2)/7.16]$$

$$= P\{-0.73 < Z < 0.67\} = \psi(0.67) + \psi(0.73) = 0.2486 + 0.2673 = 0.5159.$$

Exactly 516 candidates secure marks between 43-53 range.

7*. If X denotes the marks obtained by a student then X is $N(\mu, \sigma^2)$, where $\mu = 65.2$ and $\sigma = 5$. If p denotes the probability that a student sected at random has marks over 75, then

$$p = P(X > 75) = P(Z > 1.96) = 0.025 \qquad [Z = (X - 65.2)/5]$$

Now, the probability of getting 3 students scoring above 75 marks is a binomial probability

$$P(Y = 3) = {}^5C_3 q^2 p^3 = 10(0.25)^3 (0.975)^2$$

8*. The sizes and the upper limit of foot-length x are :

Size	x	z	Φ	d	D
2	8.75	−2.6	0.0047	0.0047	47
3	9.00	−1.6	0.0548	0.0501	501
4	9.25	−0.6	0.2743	0.2195	2195
5	9.50	0.4	0.6554	0.3811	3811
6	9.75	1.4	0.9192	0.2638	2638
7	10.00	2.4	0.9918	0.0726	726
8	10.25	3.4	0.9997	0.0079	79

In this table, $z = (x - \mu)/\sigma = 4(x - 9.40)$, $\phi = P(Z \le k)$, d stands for difference and $D = 10000 \times d$.

Notice that $\sum D_i = 997$, so that 3 women in every 10,000 requiring a size larger than 8 are not catered for.

Chapter 17 : Central Limit Theorem. Normal Approximations

Sec. 17-32. Page 534

1*. Let $X = X_1 + X_2 + \ldots + X_n$, where $X_i \sim$ Pois(3) are i.i.d. variates. Then

$$E(X) = nE(X_1) = 3n.$$

$\text{Var}(X) = n\,\text{Var}(X_1) = 3n$. Thus $S_n^* = (X - 3n)/\sqrt{3n}$ and $S_n^* \to N(0, 1)$ as $n \to \infty$.

2*. Here $E(X_i) = 1\left(\frac{1}{2}\right) + (-1)\left(\frac{1}{2}\right) = 0$, $\text{Var}(X_i) = E(X_i^2) = 1$. $S_n = X_1 + X_2 + \ldots + X_n$. Then $E(S_n) = 0$, $\text{Var}(S_n) = n$. Let

$$S^*_n = (S_n - 0)/\sqrt{n} = (X_1 + X_2 + \ldots + X_n)/\sqrt{n}. \qquad [S_n^* \text{ is standardized variate}]$$

$$\phi(t : X) = E(e^{itX}) = \tfrac{1}{2}e^{it} + \tfrac{1}{2}e^{-it} = \cos t \qquad [\phi(t) = M(it)]$$

$\therefore \quad \phi(t : S_n^*) = \phi(t/\sqrt{n} : \Sigma X_i) = [\phi_X(t/\sqrt{n})]^n = [\cos(t/\sqrt{n})]^n = [1 - (t^2/2n) + O(1/n)]^n$

where we used series expansion of $\cos(t/\sqrt{n})$. Now use Euler's limit, as $n \to \infty$.

$\lim \phi(t : S_n^*) = \lim [1 - (t^2/2n) + O(1/n)]^n = e^{-1/2t^2}$ = Ch. function of $N(0, 1)$.

It follows that $S_n^* \to N(0, 1)$ as $n \to \infty$, thus showing that C.L.T. holds for the Seq. $\langle X_n \rangle$.

3*. $P\{\overline{X} \geq \varepsilon\} \approx P\left\{\dfrac{\overline{X} - 0}{\sigma/\sqrt{n}} \geq \dfrac{\varepsilon}{\sigma/\sqrt{n}}\right\} \approx P\left\{Z \geq \dfrac{\varepsilon\sqrt{n}}{\sigma}\right\}$, (by C.L.T), $\qquad [Z \sim N(0, 1)]$

Thus $P\{\overline{X} \geq \varepsilon\} = 1 - P\{Z \leq (\varepsilon\sqrt{n}/\sigma)\} = 1 - \Phi(\varepsilon\sqrt{n}/\sigma)$.

Since $1 - \Phi(x) \sim x^{-1}(\sqrt{2n})^{-1} e^{-x^2/2}$ [§16-73, Example 9] it follows that using $x = \varepsilon\sqrt{n}/\sigma$, (1) holds.

4*. The WLLN show that $\overline{X} \xrightarrow{p} \mu$ in the sense that the $P\{\overline{X} - \mu| > \varepsilon\} \to 0$ as $n \to \infty$. By Chebyshev's inequality : $P\{\overline{X} - \mu| > \varepsilon\} \leq \sigma^2/(n\varepsilon^2)$ show that the rate of convergence is that of n^{-1}. However, C.L.T. provides a more precise estimate of the rate of convergence :

$$P\{\overline{X}_r - \mu| > \varepsilon\} = P\left\{|S_n^*| > \frac{\sqrt{n}\,\varepsilon}{\sigma}\right\} = \frac{2}{\sqrt{2\pi}}\int_k^\infty e^{-x^2/2}\,dx = \frac{2\sigma}{\sqrt{2\pi}\,\varepsilon\sqrt{n}}\exp\left(-\frac{n\varepsilon^2}{2\sigma^2}\right).\left(k = \frac{\varepsilon\sqrt{n}}{\sigma}\right)$$

where we used §16-73, Example 9 : $P(Z > k) = \varphi k/k$.

5*. Let p be the required limit. Add $-\sigma^2$ to all members in (1), then multiply by $\sqrt{n}$ to get

$$p = \lim_{n\to\infty} P\left\{-1 \leq \frac{[(X_1 - \mu)^2 - \sigma^2] + \ldots + [(X_n - \mu)^2 - \sigma^2]}{\sqrt{n}} \leq 1\right\} \qquad \ldots(2)$$

Let $Z_i = (X_i - \mu)^2 - \sigma^2$; then $\mathrm{E}(Z_i) = 0$,

$\mathrm{Var}(Z_i) = \mathrm{Var}(X_i - \mu)^2 = \mathrm{E}(X_i - \mu)^4 - [\mathrm{E}(X_i - \mu)^2]^2 = 1 + \sigma^4 - \sigma^4 = 1$.

Let $S_n = \Sigma Z_i$; then $\mathrm{E}(S_n) = 0$, $\mathrm{Var}(S_n) = n$, whence the result (2) becomes

$$p = \lim_{n\to\infty} P\{-1 \leq [S_n - \mathrm{E}(S_n)]/\sqrt{\mathrm{Var}(S_n)} \leq 1\} = \lim_{n\to\infty} P\{-1 \leq S_n^* \leq 1\} = 2\psi$$

$= 2 \times 0.3413 = 0.6826$. (Draw Figure) $\qquad$ [by C.L.T.]

6*. $\qquad M(t : T_n) = M[t : \sqrt{n}(S_n/n) - \sqrt{n}/2] = e^{-t\sqrt{n}/2} M(t/\sqrt{n} : S_n)$

$$e^{-t\sqrt{n}/2}[M(t : \sqrt{n} : X_1)]^n = e^{-t\sqrt{n}/2}\left(\frac{e^{t/\sqrt{n}} - 1}{t/\sqrt{n}}\right)^n, \qquad (\text{Put } t/\sqrt{n} = 2\theta)$$

$= \{(e^\theta - e^{-\theta})/2\theta\}^n = (\sinh\theta)^n/\theta^n = [1 + \theta^2/3! + \theta^4/5! + \ldots]^n$

$$\lim_{n\to\infty} M(t : T_n) = \lim_{n\to\infty}\left[1 + \frac{t^2}{24n} + O\left(\frac{1}{n}\right)\right]^n = \exp(t^2/24) = \exp(\sigma^2 t^2/2) \text{ (say)}.$$

Hence, by Continuity Theorem, $T_n \to N(0, \sigma^2)$ as $n \to \infty$, $[\sigma^2 = 1/12]$.

$$\lim_{n\to\infty} P\{T_n \le u\} = \int_{-\infty}^{u} \frac{e^{-x^2/2\sigma^2}\,dx}{\sqrt{2\pi}\,\sigma} = \int_{-\infty}^{u\sqrt{12}} \frac{e^{-z^2/2}}{\sqrt{2\pi}}\,dz.$$

7*. $\ln Y_n = (1/n)(\ln X_1 + \ldots + \ln X_n) = (1/n)(T_1 + T_2 + \ldots + T_n)$, $\quad [T_i = \ln X_i]$

$E(T_i) = E(\ln X_i) = \mu$ (say); $\text{Var}(T_i) = \text{Var}(\ln X_i) = \sigma^2 < \infty$, by hypothesis.

If $\quad S_n = T_1 + T_2 + \ldots + T_n$, then $E(S_n) = n\mu$, $\text{Var}(S_n) = n\sigma^2$. Now

$$S_n^* = \frac{S_n - E(S_n)}{\sqrt{\text{Var}(S_n)}} = \frac{n\ln Y_n - n\mu}{\sigma\sqrt{n}} = \frac{\sqrt{n}\ln(Y_n/e^{\mu})}{\sigma} = \frac{\ln(kY_n)^{\sqrt{n}}}{\sigma}. \quad [k = e^{-\mu}]$$

For large n, using C.L.T. $S_n^* = [\sigma^{-1}\ln(kY_n)^{\sqrt{n}}] \sim N(0,1)$. So if f is the p.d.f. of $N(0, 1)$, then

$$P\left\{\frac{\ln(kY_n)^{\sqrt{n}}}{\sigma} \le \frac{\ln t}{\sigma}\right\} = \int_{-\infty}^{(\ln t)/\sigma} f(x)\,dx \Rightarrow P\{kY_n\}^{\sqrt{n}} \le t\} = \int_{-\infty}^{(\ln t)/\sigma} f(x)\,dx.$$

So $\quad F_Z(t) = \Phi[(\ln t)/\sigma], \quad Z = (kY_n)^{\sqrt{n}}, \quad k = e^{-\mu}.$

8*. Consider an i.i.d. sequence $\langle X_n\rangle$ with common p.d.f. $f(x) = k/x^3(\ln x)^2$, $e < x < \infty$ where k is a norming constant. Being a sequence of i.i.d variates, Lindberg condition is automatically satisfied. Further

$$E(|X|^{2+\delta}) = k\int_e^{\infty} \frac{dx}{x^{1-\delta}(\ln x)^2} = k\int_e^{\infty} \frac{e^{\delta z}\,dz}{z^2}. \qquad [z = \ln x]$$

This integral is divergent as $\delta > 0$, hence $E(|X|^{2+\delta})$ is not finite. Consequently, Lyapounov condition does not hold.

9*. Let $\langle X_n\rangle$ be a sequence of independent $N(0, \sigma_n^2)$ variates where

$$\sigma_1^2 = 1, \sigma_k^2 = 2^k/4 \quad \text{for} \quad k \ge 2.$$

Let $S_n = X_1 + \ldots + X_n$, then $E(S_n) = 0$ and $B_n^2 = \text{Var}(S_n) = 1 + (1/4)[2^2 + 2^3 + \ldots + 2^n] = 2^{n-1}$.

Further, $\text{Var}(X_k/B_k) = \sigma_k^2/B_k^2 = 2^k/4 . 2^{k-1} = 1/2$, it follows that $(X_k/B_k) \sim N(0, 1/2)$

As $S_n^* = (S_n/B_n) \sim N(0,1)$ for each n [by hypothesis], the sequence $\langle X_n\rangle$ obeys C.L.T.

(i) $$\lim_{n\to\infty} \max_{1\le k\le n} \frac{\sigma_k^2}{B_n^2} = \lim_{n\to\infty}\left(\frac{2^n/4}{2^{n-1}}\right) = \frac{1}{2} \ne 0.$$

(ii) $\max_{1\le k\le n} P\{(|X_k|/B_n) \ge \varepsilon\} \ge P\{|X_n|/B_n) \ge \varepsilon\} = 1 - P\{|X_n/B_n| < \varepsilon\} > 0,$ (Non-zero)

Thus neither (*F*) nor (UAN) conditions hold. This implies that (L)-condition also does not hold. However, despite these facts, the sequence $\langle X_n\rangle$ does satisfy C.L.T.

Note. $Y = (X_n/B_n) \sim N(0, 1/2)$, so letting $Z = (Y - 0)/\sqrt{(1/2)} = \sqrt{2}Y$, we get

$P\{|Y| < \varepsilon\} = P\{|Z| < \sqrt{2}\,\varepsilon\} = P\{-\sqrt{2}\,\varepsilon < Z < \sqrt{2}\,\varepsilon\} = 2\psi(\sqrt{2}\,\varepsilon) < 1.$ $\qquad$ [use in (ii)]

Sec. 17-53. Page 540

1*. Total of 7 occurs (with two dice) with probability $p = 1/6$. Hence $X \sim$ bin (600, 1/6). Hence $\mu = np = 600/6 = 100$, $\sigma^2 = npq = 100 \times 5/6 = 225/3$. Using bin (n, p).

$$P\{90 \le X \le 110\} = \sum_{k=90}^{110} \binom{600}{k} (q)^{600-k} p^k. \qquad [p = 1/6,\ q = 5/6,\ n = 600]$$

This evaluation is not pleasing. We use C.L.T.

Write $Z = (X - np)/\sqrt{npq} = (X - 100)/\sqrt{500/6}$. Thus $X = 100 + Z\sqrt{500/6}$.

$$P\{90 \le X \le 110\} = P\{-1.095 \le Z \le 1.095\} = 2\psi\,(1.095) = 2(0.3630) = 0.0726.$$

2*. Here $Y = (X/n)$ is proportion, with $\mathsf{E}(Y) = \mathsf{E}(X/n) = np/n = 0.52$. $\quad [X \sim \text{bin}\,(n, p)]$

Var $(Y) = \sigma^2/n^2 = npq/n^2 = (0.52)\,(0.48)/n$. Now by hypothesis, using C.L.T.

$$0.01 = P\left(Y < \frac{1}{2}\right) = P\left\{\frac{Y - \mathsf{E}(Y)}{\sigma_Y} < \frac{0.50 - 0.52}{\sqrt{(0.52)\,(0.48)/n}}\right\} = P\left\{Z < \frac{-2\sqrt{n}}{\sqrt{52 \times 48}}\right\}$$

where $Z \sim N(0, 1)$. Using Tables of Normal distribution,

$$\frac{2\sqrt{n}}{\sqrt{52 \times 48}} = 2.33 \Rightarrow 4n = (2.33)^2\,(52 \times 48) \Rightarrow n = 3387.$$

3*. (a) $p = q = 1/2, n = 10$. Let $X \sim \text{bin}\,(n, p)$; $P(X = x) = {}^{10}C_x\,(1/2)^{10}$.

$$\therefore \quad P(X = 3 \text{ or } 4 \text{ or } 5) = P(X = 3) + P(X = 4) + P(X = 5)$$

$$= \left(\frac{1}{2}\right)^{10} \left[\binom{10}{3} + \binom{10}{4} + \binom{10}{3}\right] = \frac{[120 + 210 + 252]}{1024} = \frac{582}{1024} = 0.568.$$

(b) Here $np = 5$, $\sigma = \sqrt{npq} = \sqrt{2.5} = 1.581$; $Z = (X - np)/\sigma \Rightarrow X = 5 + 1.581\,Z$.

We utilize continuity correction to approximate discrete distribution :

$$P(2.5 < X < 5.5) = P(2.5 < 5 + 1.581\,Z < 5.5) = P(-1.58 < Z < 0.32)$$

$$= \psi(0.32) + \psi(1.58) = 0.1255 + 0.4429 = 0.5684.$$

There is an excellent accord between the two results.

4*. Here $p = P\{\text{wearing matching socks}\} = ({}^7C_2 + {}^5C_2)/{}^{12}C_2 = 31/66, q = 35/66$.

Now half of the time is $\frac{1}{2}(365) = 182.5$ days. Let X be the number of days she wears matching socks. Put $Z = (X - np)/\sigma$, *i.e.* $X = 365 \times (31/66) + [\sqrt{365 \times 31 \times 35/66}]\,Z$

$$= 17.1.43939 + 9.5349268\,Z. \text{ We require :}$$

$$P\{X \ge 183\} = 1 - P\{X < 183\} = 1 - P\{X \le 182 + (1/2)\}, \text{[Continuity Correction]}$$

$$= 1 - P\{171.44 + 9.535\,Z \le 182.5\} = 1 - P\,\{Z \le 11.061/9.535\}$$

$$= 1 - \Phi(1.160) = 1 - 0.8770 = 0.1230.$$

5*. Here $p = 0.30$ $q = 0.70$, $n = 100$, $x > 40$. We need $P\{X > 40\}$ which we approximate by Normal probability law using continuity correction. Here the lower limit l is given by

$$l = (x - np - \tfrac{1}{2}) / \sqrt{npq} = (40 - 30 - \tfrac{1}{2}) / \sqrt{21} = 2.07$$

$$P\{X \geq 40\} = P(Z \geq l) = P(Z \geq 2.07) = 1 - \Phi(2.07) = 1 - 0.9808 = 0.0192.$$

Thus, $P = 0.02$. Hence it is doubtful whether so many glass panes shall be smashed at this distance.

6*. Let X_i represent the weight of ith container and let $S = X_1 + X_2 + \ldots + X_{25}$; $X \sim N$ (150, 152) are independent variates, we have

$$E(S) = 25\,E(X_1) = 25 \times 150 = 3750 \text{ lb}, \text{Var}(S) = 25 \text{ Var}(X_1) = 25 \times 15^2 \Rightarrow \sigma = 75 \text{ lb}.$$

Let $Z = [S - E(S)] / \sigma_s$ i.e. $S = 3750 + 75Z$. If p is the probability of overloading the truck, then

$$p = P(S > 4000) = P(3750 + 75Z > 4000) = P(Z > 3.333) = .5 - \psi(3.33) = .5 - .4996 = .0004$$

Thus, on the average, the truck shall be over-loaded 4 times in every 10,000 occasions.

7*. (i) Let a be the 90th percentiles then $0.90 = \Phi(a_0) = \Phi[(a - 500)/100]$. (by standardization) Consulting $N(0, 1)$ tables we get $(a - 500)/100 = 1.28 \Rightarrow a = 500 + 100 \times 1.28 = 628$.

(ii) Let Q_3 be the upper quartile, its standard value being $Q'_3 = (Q_3 - 500)/100$. Then

$0.75 = \Phi(Q'_3) = \Phi[(Q_3 - 500/100]$.

Consulting $N(0, 1)$-tables we get $(Q_3 - 500/100 = 0.67 \Rightarrow Q_3 = 500 + 100 \times 0.67 = 567$.
If continuity correction is applied $Q_3 = 567.5$.
Since the lower quartile is symmetrically placed, we have

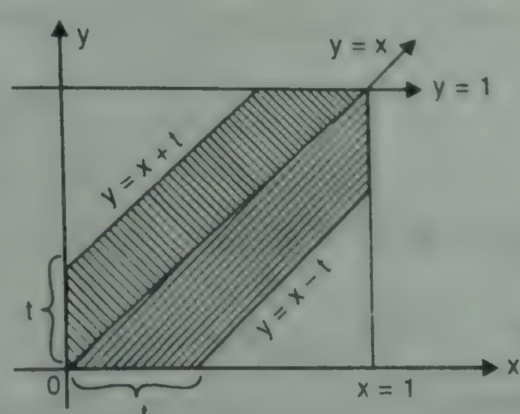
$$(500 - Q_1)/100 = 0.67 \Rightarrow Q_1 = 500 - 67 = 433.$$

If continuity correction is applied, $Q_1 = 432.5$.

Chapter 18 : Uniform Distribution. Exponential Distribution

Sec. 18-42. Page 552

1*. Let X and Y denote the arrival time of the man and the woman. Then they will meet iff, $|X - Y| \leq t$. Hence the probability p is given by



$$p = P\{|X - Y| \leq t\} = \frac{\text{Measure of shaded area}}{\text{Measure of total area}}$$

$$= 1 - (1 - t)^2 = t(2 - t), \text{ hours}.$$

When $t = 10$ min, $p = 11/36$.

2*. Since $X \sim$ Unif $(-b, b)$, its p.d.f. is $f(x) = 1/2b$, $-b < x < b$; $f(x) = 0$, otherwise

So $\quad 3/4 = P(|X|>2) = 1 - P(|X| \le 2) = 1 - P(-2 \le X \le 2) = 1 - 2\int_0^2 \frac{dx}{2b}.$

Thus : $\quad 2/b = 1/4 \quad \Rightarrow b = 8.$

3*. Since $X \sim$ Unif (1, 2) ; $f(x) = 1,\ 1 \le x \le 2$; $\mu_X = (1+2)/2 = 3/2$,

$$\therefore \qquad \frac{1}{4} = P\left\{X > z + \frac{3}{2}\right\} = \int_{z+(3/2)}^{2} dx = 2 - \left(z + \frac{3}{2}\right).$$

Thus $\quad z = 1/4.$

4*. Here $\mu_X = (a+b)/2 = 1$, $\sigma^2_X = (b-a)^2/12 = 4/3$. Chebyshev's inequality

$$P\{X - \mu| \ge c\} \le \sigma^2/c^2 \quad \Rightarrow P\{|X - \mu| \ge 2\sigma\} \le 1/4.$$

Thus, the upper bound to the given probability is 1/4.

To obtain the exact value of probability, we not that $f(x) = 1/4$, so

$$\begin{aligned} P\{|X-1| \ge 4/\sqrt{3}\} &= 1 - P\{|X-1| < 2.61\} = 1 - P\{-1.61 < X < 3.61\} \\ &= 1 - P\{-1 < X < 3\} = 1 - 1 = 0. \qquad [\because\ x \in (-1, 3)] \end{aligned}$$

5*. We need $p = P(L \text{ or } M) = P(L \cup M) = 1 - P(L' \cap M')$, ...(1)

$$\begin{aligned} P\{L' \cap M'\} &= P\{\min(X_1, ..., X_n) > -b; \max(X_1, ..., X_n) < a\} = P\{-b < X_j < a,\ \forall\ j\} \\ &= \prod_{j=1}^{n} P(-b < X_j < a) = [(a+b)/2c]^n, \qquad (1 \le j \le n) \end{aligned}$$

since $P(-b < X_j < a) = \int_{-b}^{a} \frac{dx}{2c} = \frac{a+b}{2c}$. Substituting in (1) gives $p = 1 - [(a+b)/2c]^n$.

6*. Solving the given quadratic equation we get $x = \frac{1}{2}[a \pm \sqrt{a^2 - 4b}]$. Thus, the roots are imaginary if $b > (a^2/4)$. We find this probability :

$$P\left(b > \frac{a^2}{4}\right) = \int_0^6 \int_{a^2/4}^{9} \frac{1}{54}\, da\, db = \int_0^6 \left(9 - \frac{a^2}{4}\right)\frac{da}{54} = \frac{2}{3}. \qquad \left[\because\ f(a, b) = f(a)\, f(b) = \frac{1}{6}\cdot\frac{1}{9}\right]$$

$$\therefore \qquad P(b \le a^2/4) = P\ \{\text{roots are real}\} = 1/3.$$

7*. The Ch. Function of $X_i \sim$ Unif $\left(-\frac{1}{2}, \frac{1}{2}\right)$ is $\phi(t : X_i) = \int_{-1/2}^{1/2} e^{itx} dx = \dfrac{2\sin\left(\frac{1}{2}t\right)}{t}$.

Since X_1, X_2, X_3 are i.i.d. it follows that $\phi(t : X_1 + X_2 + X_3) = [\phi(t : X_1)]^3$

$$\therefore \qquad \phi(t : X) = \left[\frac{2\sin\left(\frac{1}{2}t\right)}{t}\right]^3 = \left[1 - \frac{t^2}{24} + \frac{t^4}{1920} - ...\right]^3 = 1 - \frac{t^2}{8} + \frac{13}{1920}t^4 + ...$$

Thus, $E(X^4)$ = Coeff. of $(it)^4\ ! = \dfrac{4!\,13}{1920} = \dfrac{13}{80}.$

8*. (i) $F(x) = \int_0^x 6t(1-t)\,dt = 3x^2 - 2x^3$; $G(x) = \int_0^y 3(1 - \sqrt{t})\,dt = 3y - 2y^{3/2}$.

Setting $F(x) = G(y)$, since both are Unif (0, 1) variates, we get $3y - 2y^{3/2} = 3x^2 - 2x^3$

$\Rightarrow\ y = x^2.$

(ii) $$M(t:Z) = E(e^{tZ}) = \int_0^1 e^{t(3x^2-2x^3)}.(6x-6x^2)\,dx$$

$$= \int_0^1 e^{tz}\,dz = \left(\frac{e^t-1}{t}\right)$$

This shows that $Z \sim$ Unif (0, 1).

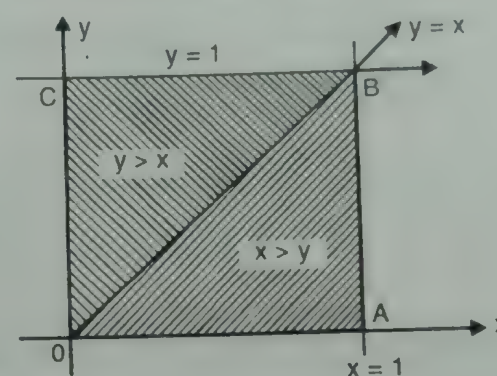
9*. (i) $F(x) = \int_0^x 2t\,e^{-t^2}dt = 1-e^{-x^2}$. Choose a random value $y \in (0, 1)$ and put

$$y_j = 1-\exp(-x_j^2) \Rightarrow x_j = -[\ln(1-y_j)]^{1/2};\ j = 1, 2, 3, \ldots, n.$$

(ii) $$F(x) = \int_a^x \frac{ra^r}{t^{r+1}}dt = 1-\left(\frac{a}{x}\right)^r.$$

Choose a random value $y \in (0, 1)$ and put $y_j = 1-(a/x_j)^r \Rightarrow x_j = a/(1-y_j)^{1/r}$.

10*. Here domain of $Z = X - Y$ is the square $OABC$



$$\phi(t:|X-Y|) = \int_0^1\int_0^1 e^{it|x-y|}\,dx\,dy$$

$$= 2\int_0^1\int_0^x e^{it(x-y)}\,dx\,dy$$

$$= 2\int_0^1 e^{itx}\,dx\int_0^x e^{-ity}\,dy$$

$$= \frac{2}{it}\int_0^1 e^{itx}(1-e^{itx})\,dx$$

$$= \frac{2}{it}\left[\frac{e^{it}-1}{it}-1\right] = 2(1+it-e^{it})\,t^{-2}.$$

To find moments, we expand exponential and get

$$\phi(t:|Z|) = -\frac{2}{t^2}\sum_{n=2}^{\infty}\frac{(it)^n}{n!} = -2\sum_{n=0}^{\infty}\frac{i^{n+2}.t^n}{(n+2)!} = 2\sum_{n=0}^{\infty}\frac{(it)^n}{n!}\cdot\frac{1}{(n+2)(n+1)}.$$

$$\mu_n' = E(|X-Y|)^n = \text{Coeff. of } (it)^n/n! = 2/(n+1)(n+2).$$

11*. Let $u = xy$, $v = x/y$, and invert them to get $x = (uv)^{1/2}$, $y = (u/v)^{1/2}$.

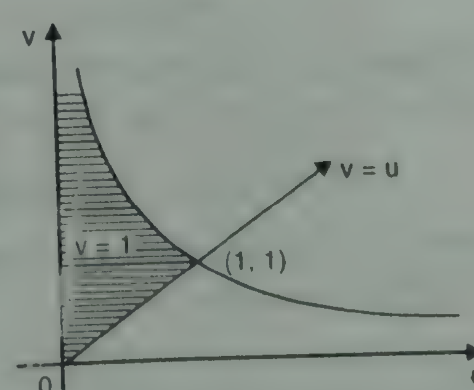
$$J = \frac{\partial(x,y)}{\partial(u,v)} = \begin{vmatrix} \frac{1}{2}(v/u)^{1/2} & \frac{1}{2}(u/v)^{1/2} \\ \frac{1}{2}(uv)^{-1/2} & -\frac{1}{2}u^{1/2}v^{-3/2} \end{vmatrix} = -\frac{1}{2}\frac{1}{v}.$$ Thus, $dx\,dy = \frac{1}{2}v^{-1}\,du\,dv$.

The region in the u-v plane is trivially obvious, as

$0 \le x \le 1,\ 0 \le y \le 1 \Rightarrow 0 \le \sqrt{uv} \le 1,\ 0 \le \sqrt{u/v} \le 1.$

i.e. $0 \le uv \le 1,\ 0 \le u \le v$. [Shaded part in Fig.]

Since X and Y are indep.; their joint elemental p.d.f. is



$$dP_1(x, y) = f_1(x)\,f_2(y)\,dx\,dy = dx\,dy$$

$$\therefore \quad dP_2(u, v) = \tfrac{1}{2}v^{-1}\,du\,dv. \qquad \ldots(1)$$

To find the p.d.f. of U, we integrate out v from Eq. (1); thus

$$dP_3(u) = \frac{1}{2}du\int_u^{1/u}\frac{1}{v}dv = -(\ln u)\,du,\quad 0<u<1,\quad \Rightarrow g_1(u) = -\ln u,\ 0<u\le 1.$$

To find the p.d.f. of V, we integrate out u from Eq. (1) ; thus

$$\left.\begin{aligned} dP_4(v) &= \frac{dv}{2v}\int_0^v du = \frac{1}{2}dv; && 0<v<1 \\ &= \frac{dv}{2v}\int_0^{v^{-1}} du = \frac{2}{2v^2}dv, && 1<v<\infty \end{aligned}\right] \Rightarrow g_2(v) = \begin{cases}\frac{1}{2}, & 0\le v<1 \\ \frac{1}{2}v^{-2}, & 1<v<\infty.\end{cases}$$

12*. If $X_i \sim$ Unif (0, 1) then p.d.f. of $S = X_1 + X_2 + X_3$, is [§ 18-41(3)]

$$g(u) = \tfrac{1}{2}u^2, 0<u<1;\ g(u) = -\tfrac{1}{2}(2u^2-6u+3),\ 1<u<2,\ g(u) = \tfrac{1}{2}(u-3)^2, 2\le u<3.$$

Now $W = 2S - 3$ so $S = (W+3)/2$; hence

$$g(u)\,du = (\tfrac{1}{2})g[(w+3)/2]\,dw \Rightarrow f(w) = \tfrac{1}{2}g[(w+3)/2].$$ This yields (A)

The c.d.f. of W is given by

$$F(w) = \begin{cases}(1/48)(w+3)^3, & -3\le w\le -1 \\ (1/24)(9w-w^3)+(1/2), & -1\le w\le 1 \\ (1/48)(w-3)^3+(44/88), & 1\le w\le 3\end{cases}$$

Using Normal tables $[\Phi(w)]$ we get

w	:	0	0.5	1.0	1.5	2.0
$F(w)$	:	0.50	0.68	0.83	0.93	0.98
$\Phi(w)$	:	0.50	0.69	0.84	0.93	0.98

Comparison reveals that $2W - 3$ is really an excellent approx. for $N(0, 1)$. [Compare §16-61]

Sec. 18-82. Page 564

1*. ***Counter Examples*** : (a) Let X and Y be discrete with p.m.f.

$$P(X=0) = P(Y=0) = 3/5,\quad P(X=1) = P(Y=1) = 2/5.$$

Then mode (X) = mode $(Y) = 0$. Now let $Z = X + Y$. Using indep. (X, Y) we have

$$\begin{aligned} P(Z=0) &= P(X=0, Y=0) = 9/25 \\ P(Z=1) &= P(X=0, Y=1) + P(X=1, Y=0) = 6/25 + 6/25 = 12/25. \\ P(Z=2) &= P(X=1, Y=1) = 4/25. \end{aligned}$$

Obviously $Z = 1$ is the modal value of Z and mode (X) + mode $(Y) = 0 \ne 1$ = mode $(X + Y)$.

(b) Let X and Y be i.i.d. Expo (1) variates ; so that $F(t) = 1 - e^{-t}$

$$F(t) = \tfrac{1}{2} \Rightarrow t = \ln 2$$ and thus med (X) + med $(Y) = \ln 2 + \ln 2 = \ln_e 4$.

Now $Z = (X + Y) \sim$ gam (2, 1), so $f(z) = ze^{-z}$. [§8-18 (3)]

If med $Z = m$, then

$$\frac{1}{2} = \int_0^m ze^{-z}\,dz = \left[(z+1)e^{-z}\right]_m^0 = 1-(m+1)e^{-m}.$$

Thus $2(m+1) = e^m$. If we try $m = \ln 4$, we get $\frac{1}{2} = \ln 2$, an obviously wrong result. Hence med X + med $Y \neq$ med $(X + Y)$.

2*. Here $f(x) = \lambda e^{-\lambda x}, 0 < x < \infty, \lambda = 1/3000$. Let Y denote the daily consumption of milk, then $X = Y - 20{,}000$ is the exponential variate of interest. The stock shall be insufficient on any day if consumption on that day exceeds 35000 gallons. Hence the probability p for this event is given by

$$p = P\{Y > 35000\} = P\{X > 15000\} = \int_{15000}^{\infty} \lambda e^{-\lambda x}\, dx = \int_5^{\infty} e^{-z}\, dz = e^{-5}, \quad (z = \lambda x)$$

$\therefore \quad P\{\text{Stock insufficient on both the days}\} = {}^2C_2\, q^{2-2}\, p^2 = p^2 = e^{-10}$.

3*. Here $f(x) = \lambda e^{-\lambda x}, x \geq 0$, and so $F(k) = P(X \leq k) = 1 - e^{-\lambda k}, P(X > k) = 1 - P(X \leq k) = e^{-\lambda k}$.

$\therefore \quad [P(X > k)/P(X \leq k)] = e^{-\lambda k}/(1 - e^{-\lambda k}) = a$, by hypothesis.

Thus $e^{+\lambda k} = (1+a)/a \Rightarrow \lambda k = \ln(1 + a^{-1}) \Rightarrow x_a = \lambda^{-1} \ln(1 + a^{-1})$. $\quad (k = x^k)$

4*. $P(X \leq 1) = P(X > 1) = 1 - P(X \leq 1) \Rightarrow P(X \leq 1) = \frac{1}{2}$ (by transfer). Also $F(k) = 1 - e^{-\lambda k}$.

Thus $1 - e^{-\lambda} = \frac{1}{2}$ i.e. $e^{-\lambda} = \frac{1}{2} \Rightarrow \lambda = \ln_e 2$. So Var $(X) = 1/\lambda^2 = 1/(\ln 2)^2$.

5*. $M(t : Y) = \mathsf{E}[\exp(-t/\lambda) \ln F(X)] = \mathsf{E}[\exp \ln [F(X)]^{-t/\lambda} = \mathsf{E}[F(X)]^{-t/\lambda}$

$$= \int_0^1 (z)^{-t/\lambda}\, dz = [1 - (t/\lambda)]^{-1}. \qquad [Z = F(X) \sim U(0, 1)].$$

Thus, $\;Y \sim$ Expo (λ), i.e. $f(y) = \lambda e^{-\lambda y}, y \geq 0$.

6*. $F_T(t) = P\{T \leq t\} = P\{Y \leq tX\} = \int_{x=0}^{\infty} \int_{y=0}^{tx} \lambda^2 e^{-\lambda x} . e^{-\lambda y} dx\, dy = \lambda \int_0^{\infty} e^{-\lambda x} \left[e^{-\lambda y}\right]_{tx}^{0} dx$

$$= \lambda \int_0^{\infty} e^{-\lambda x} (1 - e^{-\lambda tx})\, dx = \int_0^{\infty} \lambda e^{-\lambda x}\, dx - \lambda \int_0^{\infty} x^{1-1} e^{-\lambda(1+t)x} dx$$

$$= 1 - \{\Gamma(1)/(1+t)\} = 1 - 1/(1+t)$$

$\therefore \quad f(t) = (dF/dt) = 1/(1+t)^2, 0 < t < \infty$.

Note. $\mathsf{E}(T) = \int_0^{\infty} t\, dt/(1+t)^2 \to \infty$; i.e. $\mathsf{E}(T)$ does not exist.

7*. Here $\lambda = 1/n$, so that $f(x) = n^{-1} e^{-x/n}, \quad x \geq 0$.

Let $X_i \sim$ Expo (λ), $1 \leq i \leq k$, be independent variates. We want :

$p = P\{\text{First failure occurs within } m \text{ months}\} = P\{\text{Atleast one } X_j \leq m\} = 1 - P(X_j > m, \forall j)$

$= 1 - P(X_1 > m, X_2 > m, \ldots, X_k > m) = 1 - [P(X_1 > m)]^k$, $\qquad (\because X_j$ are i.i.d.)

Since $P(X_i > m) = 1 - F_i(m) = e^{-m/n}$, hence $p = 1 - (e^{-m/n})^k = 1 - e^{-mk/n}$.

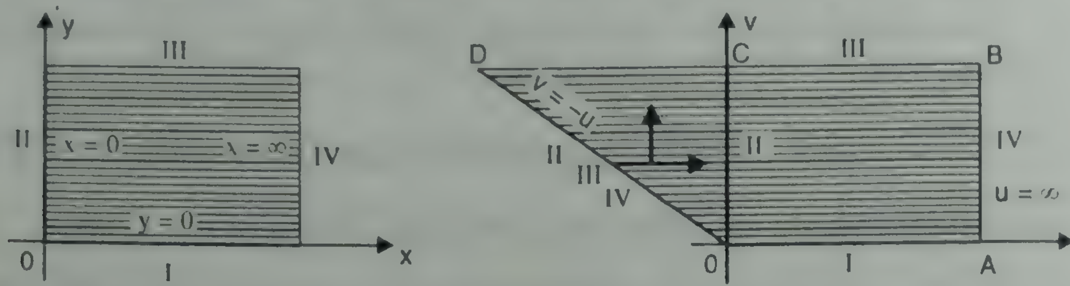
8*. (a) the expected length of an interval is $\mathsf{E}(T) = 1/\lambda$ = mean Expo (λ).

$$P(T > 1/\lambda) = \int_{1/\lambda}^{\infty} \lambda e^{-\lambda t}\, dt = e^{-1} = 0.37.$$

(b) $\quad P\{T < (1/\lambda)\} = 1 \times P\{T \geq (1/\lambda)\} = 0.63 = p$, say.

The required probability p_1 is now bin $(5, p)$; here $n = 5$, $x = 2$, $p = 0.63$, $q = 0.37$ so that $p_1 = {}^5C_2\, q^3 p^2 = 10 \times (0.37)^3 \times (0.63)^2 = 0.2010$.

9*. Let $u = x - y$, $v = y$ so that $x = u + v$, $y = v$. Also $|\,\partial(x, y)/\partial(u, v)\,| = 1$, hence $dxdy = dudv$. The domains of definition of (x, y) and (u, v) are shown.



I. $y = 0,\ 0 \le x < \infty;\ \Rightarrow v = 0,\ 0 \le u < \infty$.   II. $x = 0,\ 0 \le y < \infty;\ \Rightarrow v = -u,\ 0 \le v < \infty$.

III. $0 \le x < \infty,\ y = \infty;\ \Rightarrow 0 \le u + v < \infty,\ v = \infty$. IV. $0 \le y < \infty,\ x = \infty;\ \Rightarrow u + v < \infty,\ 0 \le v < \infty$.

The joint elemental p.d.f. of indep. X and Y is $dP_1(x, y) = \alpha\beta e^{-\alpha x}\, e^{-\beta y}\, dx\, dy$;

In terms of new variates : $dP_2(u, v) = \alpha\beta e^{-\alpha u}\, e^{-(\alpha+\beta)v}\, du\, dv$. Integrating out v we obtain

$$dP(u) = \alpha\beta e^{-\alpha u}\, du . \begin{cases} \int_{-u}^{\infty} e^{-(\alpha+\beta)v}\, dv, & u < 0 \\ \int_0^{\infty} e^{-(\alpha+\beta)v}\, dv, & u > 0 \end{cases}. \qquad \text{Thus } dP(u) = \begin{cases} [\alpha/(\alpha+\beta)]e^{\beta u}, & u < 0 \\ [\alpha/(\alpha+\beta)]e^{-\alpha u}, & u \ge 0 \end{cases}.$$

10*. Probability Integral Transform shows that c.d.f. $F_X(x)$ and $F_Y(y)$ are $U(0, 1)$ variates. Now

$$F_X(x) = \int_0^x \lambda e^{-\lambda t}\, dt = 1 - e^{-\lambda x},\ x \ge 0;\ F_X(x) = 0, \text{ otherwise}$$

$$F_Y(y) = \int_0^y \tfrac{1}{2} t^{-1/2}\, dt = y^{1/2},\ 0 \le y \le 1;\ F_Y(y) = 0, \text{ otherwise.}$$

Thus $Z = 1 - e^{-\lambda x}$ is $U(0, 1)$, $W = \sqrt{Y}$ is $U(0, 1)$. Changing W to Z provides the relation $y = [1 - e^{-\lambda x}]^2$. This means, if X has prob. law f, then $[1 - e^{-\lambda X}]^2$, shall have the p.d.f. "g".

11*. For a mixed variate X,

$$M(t:X) = \mathrm{E}(e^{tX}) = \int_0^{\infty} \frac{e^{tx}.e^{-x/\lambda}}{2\lambda}\, dx + \sum_{x=0}^{\infty} \frac{e^{-\lambda}\lambda^x}{x!}\frac{e^{tx}}{2} = \tfrac{1}{2}[\text{m.g.f. of expo } (\lambda^{-1})] + \tfrac{1}{2}[\text{m.g.f. of Pois } (\lambda)]$$

$$= \tfrac{1}{2}(1 - \lambda t)^{-1} + \tfrac{1}{2}\exp[\lambda(e^t - 1)].$$

We may expand these exponential functions to obtain

$$M(t:X) = \tfrac{1}{2}(1 + \lambda t + \lambda^2 t^2 + \ldots) + \tfrac{1}{2}\exp\left[\lambda\left(t + \frac{t^2}{2!} + \frac{t^3}{3!} + \ldots\right)\right]$$

$$= \tfrac{1}{2}(1 + \lambda t + \lambda^2 t^2 + \ldots) + \tfrac{1}{2}\left\{1 + \lambda\left(t + \frac{t^2}{2!} + \ldots\right) + \frac{\lambda^2}{2!}\left(t + \frac{t^2}{2!} + \ldots\right)^2 + \ldots\right\}$$

Thus μ = Coeff. of $t = \lambda$, μ'_2 = Coeff. of $t^2/2! = (3/2)\lambda^2 + \tfrac{1}{2}\lambda \Rightarrow \text{Var}(X) = \tfrac{1}{2}(\lambda + \lambda^2)$.

Note. Using moments of expo (λ^{-1}) and Pois (λ) we can directly evaluate $\mathsf{E}(X)$ and Var (X).

Thus $\mathsf{E}(X) = \frac{1}{2}\int_0^\infty \frac{xe^{-x/\lambda}}{\lambda}\,dx + \frac{1}{2}\sum_{x=0}^{\infty}\frac{x.e^{-\lambda}\lambda^x}{x!} = \frac{1}{2}(\lambda+\lambda) = \lambda.$

$$\mathsf{E}(X^2) = \tfrac{1}{2}[2\lambda^2 + (\lambda+\lambda^2)] = (3/2)\lambda^2 + \lambda/2, \text{ etc.}$$

12*. Let $W = \max\{Y_1, Y_2, \ldots, Y_N\}$. Firstly, we determine W. Let $F(w)$ be the c.d.f. of W. Now

$$F(w) = P\{W \le w\} = \sum_{n=1}^{\infty} P\{W \le w, N = n\} = \sum_{n=1}^{\infty} P\{Y_1 \le w, Y_2 \le w, \ldots, Y_n < w, N = n\}$$

$$= \sum_{n=1}^{\infty}[P(Y_i \le w)]^n\, P(N=n) = \sum_{n=1}^{\infty}[G_i(w)]^n\, p_n \qquad [\because\ Y_i \text{ is indep. of } N] \quad \ldots(1)$$

where $p_n = P\{N = n\}$ and $G_i(w) = P\,\{Y_i \le w\}$, is the c.d.f. of $Y_i \sim U(0, 1)$. As such the functions $G_i(w)$ are given by

$$G_i(w) = 0,\ w < 0, \quad G_i(w) = w,\ 0 \le w < 1, \quad G_i(w) = 1,\ w > 1. \quad [\text{c.d.f. of } U(0, 1)]$$

Substitutions into (1) now provide

$$F(w) = 0,\ w < 0; \quad F(w) = \sum w^n p_n,\ 0 \le w \le 1; \quad F(w) = \sum (1)^n p^n,\ w > 1. \quad \ldots(2)$$

Now $\sum_{n=1}^{\infty} p_n = \frac{1}{e-1}\sum_{n=1}^{\infty}\frac{1}{n!} = \frac{1}{e-1}(e-1) = 1; \qquad \sum_{n=0}^{\infty} w^n p_n = \frac{1}{e-1}\sum_{n=0}^{\infty}\frac{w^n}{n!} = \frac{e^w - 1}{e-1}.$

$\therefore \quad F(w) = 0,\ \ w < 0;\ F(w) = (e^w - 1)/(e-1),\ 0 \le w < 1;\ F(w) = 1,\ w > 1. \quad$ [by (2)]

So $\qquad f(w) = F'(w) = e^w/(e-1),\ 0 \le w \le 1.$

$$\phi(t : Z) = \phi(t : X - W) = \phi(t : X)\,\phi(-t : W) \qquad (\because\ X \text{ is indep. of } W) \quad \ldots(3)$$

We now calculate the Ch. Functions separately.

$$\phi(t : X) = \mathsf{E}(e^{itx}) = \sum_{m=1}^{\infty}(e-1)\,e^{-m}e^{itm} = (e-1)\sum_{m=1}^{\infty}T^m = \frac{(e-1)T}{1-T}. \qquad [T = e^{it-1}]$$

$$\phi(-t : W) = \int_0^1 e^{-itw}\,\frac{e^w\,dw}{e-1} = \int_0^1 \frac{e^{w(1-it)}}{e-1}\,dw = \frac{(e^{1-it} - 1)}{(e-1)(1-it)} = \frac{(T^{-1} - 1)}{(e-1)(1-it)}.$$

$$\therefore \quad \phi(t : Z) = \frac{(e-1)T}{1-T}\cdot\frac{1-T}{T}\,\frac{1}{(e-1)(1-it)} = \frac{1}{1-it} \qquad [\text{by (3)}]$$

This shows that $Z \sim$ Expo (1).

13*. $p_1 = \int_0^{1/\lambda}\lambda e^{-\lambda x}\,dx = \frac{e-1}{e},\ p_2 = \int_{1/\lambda}^{2/\lambda}\lambda e^{-\lambda x}\,dx = \frac{e-1}{e^2},\ p_3 = 1 - p_1 - p_2 = \frac{1}{e^2}.$

Recall : $(X, Y) \sim$ Trin $(3, p_1, p_2, p_3)$. Since $q_1 = 1 - p_1 = 1/e$, $q_2 = 1 - p_2 = (e^2 - e + 1)/e^2$ the known result [§ 15-34] gives

$$\text{Corr}(X, Y) = -\sqrt{p_1p_2/q_1q_2} = -(e-1)/\sqrt{(e^2 - e + 1)}.$$

14*. Cov $(X, Y) = 0$ need not provide independence of X and Y. Thus, it is impossible to determine the joint density of X and Y from the given data.

Note. Where X and Y independent, then $f(x, y) = f_X(x)\,f_Y(y) = ae^{-ax}\,be^{-by},\ x \ge 0,\ y \ge 0$!

Chapter 19 : More Continuous Distributions

Sec. 19-15. Page 571

1*. If X is gam (α, λ), then $\mathsf{E}(X) = \alpha/\lambda$, Mode $= (\alpha - 1)/\lambda$, hence $a = \alpha/\lambda$, $b = (\alpha - 1)/\lambda$; these give $(1/\lambda) = a - b$. Now Var $(X) = \alpha/\lambda^2 = (\alpha/\lambda)\,(1/\lambda) = a\,(a - b)$.

[For mode of $\Gamma(\alpha;\lambda)$ see Ex. 1, Set Ex. 19(a).

2*. With given μ_n', we determine the m.g.f. of X. Thus

$$M(t : X) = \sum_{n=0}^{\infty} \mu_n' \cdot \frac{t^n}{n!} = \sum_{n=0}^{\infty} \frac{(n+k)!}{n!\;k!}(kt)^n = \sum_{n=0}^{\infty}\binom{n+k}{n}(kt)^n = (1-kt)^{-(k+1)}.$$

where we used Neg-bin Expansion. This is the m.g.f. of gam $(k + 1,\ 1/k)$, hence

$$f(x) = (k)^{-(k+1)}\, e^{-x/k}\, x^k / \Gamma(k+1),\quad 0 \le x < \infty.$$

3*. Let $Z = (X - \mu)/\sigma \sim N(0, 1)$. Then $Y = Z^2/2\lambda$, and so

$$M_Y(t) = \mathsf{E}(e^{tZ^2/2\lambda}) = \int_{-\infty}^{\infty} \frac{e^{-\frac{1}{2}z^2}}{\sqrt{2\pi}} . e^{tz^2/2\lambda}\, dz = \int_{-\infty}^{\infty} e^{-\frac{1}{2}z^2(1-t/\lambda)}\,\frac{dz}{\sqrt{2\pi}}$$

$$= \frac{1}{\sqrt{(1-t/\lambda)}}\int_{-\infty}^{\infty} \frac{e^{-\frac{1}{2}u^2}}{\sqrt{2\pi}}\,du = \left(1 - \frac{t}{\lambda}\right)^{-\frac{1}{2}},\quad \left[u \equiv z\sqrt{1-\frac{t}{\lambda}}\right].$$

Thus $Y \sim$ gam $(1/2, \lambda)$, hence Var $(Y) = \alpha/\lambda^2 = 1/2\lambda^2$.

4*. If X is gam (r, λ) then

$$\mathsf{E}(\sqrt{X}) = \int_0^{\infty}\left(\frac{\lambda^r e^{-\lambda x} x^{r-1}}{\Gamma(r)}\right)x^{1/2}\,dx = \frac{\lambda^r}{\Gamma(r)}\int_0^{\infty} e^{-\lambda x}.x^{r+1/2-1}\,dx = \frac{\lambda^r}{\Gamma(r)}\cdot\frac{\Gamma(r+\frac{1}{2})}{(\lambda)^{r+1/2}} = \frac{\Gamma(r+\frac{1}{2})}{\sqrt{\lambda}\,\Gamma(r)}.$$

Now if Y is $N(\mu, \sigma^2)$, then $Z = (Y - \mu)^2/2s^2$ is gam $(\frac{1}{2}, 1)$.

$$\therefore \qquad \mathsf{E}(|Y - \mu|) = \sigma\sqrt{2}\,\mathsf{E}(\sqrt{Z}) = \sigma\sqrt{2}\,\{\Gamma(\tfrac{1}{2}+\tfrac{1}{2})/\Gamma(\tfrac{1}{2})\} = \sigma(2/\pi)^{1/2}.$$

5*. The joint p.d.f. of X_1 and X_2 is

$$f(x_1, x_2)\,dx_1\,dx_2 = \lambda^2 e^{-\lambda(x_1+x_2)}\,dx_1\,dx_2,\ 0 < x_1, x_2 < \infty.$$

Also $\qquad x_1 = y_1 y_2,\ x_2 = y_1(1 - y_2)$ so that

$$\frac{\partial(x_1, x_2)}{\partial(y_1, y_2)} = \begin{vmatrix} \partial x_1/\partial y_1 & \partial x_2/\partial y_1 \\ \partial x_1/\partial y_2 & \partial x_2/\partial y_2 \end{vmatrix} = \begin{vmatrix} y_2 & 1-y_2 \\ y_1 & -y_1 \end{vmatrix} = -y_1.$$

Thus, $dx_1\,dx_2 = |\,y_1\,|\,dy_1\,dy_2$, and the probability differential element for y_1, y_2 is

$$g(y_1, y_2) = \lambda^2 e^{-\lambda y_1}.y_1,\ 0 < y_1 < \infty,\ 0 < y_2 < 1.\ [g(y_1, y_2) = g_1(y_1).g_2(y_2)]$$

This shows that Y_1 is gam $(2, \lambda)$ and Y_2 is $U(0, 1)$ and that they are indep. distributed.

6*. Since X is gam $(m, 1/a)$ its m.g.f. is $M(t : X) = (1 - at)^{-m}$ ...(1)

$\therefore \quad M(t:\overline{X}) = M(t/n: X_1 + X_2 + ... + X_n) = [M(t/n: X_i)]^n = [1-(at/n)]^{-mn}$, [by (1)]

This shows that $\overline{X} \sim \text{gam}(mn, n/a)$. Hence $Y = \overline{X}$ has

$$f(y) = (n/a)^{mn} e^{-nz/a}(z)^{mn-1}/\Gamma(mn), \quad 0<z<\infty.$$

7*. The m.g.f. of $X \sim$ Expo (λ) is $M(t : X) = (1 - t/\lambda)^{-1}$.

$\therefore\ M(t:n\lambda\overline{X}) = M(\lambda t: X_1 + X_2 + ... + X_n) = [M(\lambda t: X_i)]^n = (1-t)^{-n}$ [$\because$ X_i are i.i.d.] ...(1)

This shows that $n\lambda\overline{X} \sim \text{gam}(n, 1)$. Now consider c.g.f.

$$K(t) = \ln M(t:n\lambda\overline{X}) = -n\ln(1-t) = n\Sigma(t^r/r), \quad 1 \le r < \infty$$

$$\therefore \quad \kappa_r = (r-1)!\,n; \qquad \kappa_1 = n,\ \kappa_2 = n.$$

So
$$E(n\lambda\overline{X}) = n, \text{Var}(n\lambda\overline{X}) = n \Rightarrow E(\overline{X}) = 1/\lambda$$

$$\text{Var}(\overline{X}) = 1/n\lambda^2 \Rightarrow \text{S.E.}(\overline{X}) = 1/\lambda\sqrt{n}.$$

8*. Here $X \sim \text{gam}_\theta\ (\alpha, \lambda)$. Now

$$M(t:X) = E(e^{tX}) = \frac{\lambda^\alpha}{\Gamma(\alpha)}\int_0^\infty e^{tx}\,.\,e^{-\lambda(x-\theta)}\,(x-\theta)^{\alpha-1}\,dx \qquad [\text{Put } x - \theta = y]$$

$$= \frac{\lambda^\alpha e^{t\theta}}{\Gamma(a)}\int_0^\infty e^{-(\lambda-t)y}\,y^{\alpha-1}\,dy = \frac{\lambda^\alpha e^{t\theta}}{\Gamma(\alpha)}\cdot\frac{\Gamma(\alpha)}{(\lambda-t)^\alpha}$$

$$= e^{\theta t}[1-(t/\lambda)]^{-\alpha}, \ (t<\lambda). \qquad ...(1)$$

Observe that if $X \sim \text{Expo}_\theta(\lambda)$, then

$$M(t) = e^{\theta t}[1-(t/\lambda)]^{-1}. \qquad ...(2)$$

We now determine the m.g.f. of $\overline{X}$ as under :

$$M(t:\overline{X}) = M[t:(X_1 + ... + X_n)/n] = M[t/n: X_1 + ... + X_n] = (M[t/n: X_1])^n$$

$$= \{e^{\theta(t/n)}[1-(t/n\lambda)]^{-1}\}^n, = e^{\theta t}[1-(t/\lambda n)]^{-n}. \quad [\text{by } (2)] \qquad ...(3)$$

From (1) and (3) we conclude that $\overline{X} \sim \text{gam}_\theta(n, n\lambda)$, hence the p.d.f. of $\overline{X}$ is

$$f(y) = (\lambda n)^n e^{-\lambda n(y-\theta)}.(y-\theta)^{n-1}/\Gamma(n), \ y \ge \theta, \ \lambda > 0.$$

9*. The Chernoff Bound valid for any variate X is $P(X \ge c\} \le \min\ e^{-ct}\, M(t:X), [0 \le t \le h]$.
Since X is gam $(\alpha\ ;\ \lambda)$, $M(t : X) = [1 - (t/\lambda)]^{-\alpha}$ and this bound provides

$$P\{X \ge 2\alpha/\lambda\} \le \min e^{-2\alpha t/\lambda}.\ \{1-(t/\lambda)]^{-\alpha} \qquad ...(1)$$

Let $w = [1-(t/\lambda)]^{-\alpha}\ .\ e^{-2\alpha t/\lambda}$; then

$$\frac{dw}{dt} = \frac{\alpha w}{\lambda}\left[\frac{1}{1-(t/\lambda)} - 2\right]; \frac{d^2w}{dt^2} = \frac{\alpha w}{\lambda^2}\left(1-\frac{t}{\lambda}\right)^{-2} + \frac{dw}{dt}\frac{\alpha}{\lambda}\left[\left(1-\frac{t}{\lambda}\right)^{-1} - 2\right].$$

Put $dw/dt = 0$ to get $t = \lambda/2$, and then $(d^2w)/(dt^2) > 0$ at $t = 1/2\lambda$.
Thus, w is min. at $t = \lambda/2$ and min $w = (2/e)^{\alpha}$. Substituting this into (1), the stated result follows.

Sec. 19-25. Page 577

1*. With the notation of § 19-24, we have $m = 6$, $M = 12$, $h = 12 - 6 = 6$, $\mu = 8$.

$$\therefore \quad \mu = m + (ah/c) \Rightarrow 8 = 6 + [6a/(a+b)] \Rightarrow b = 2a.$$

Substituting into the expression for variation (C.V. $= \sigma/\mu$) we get

$$(0.1 \times 8)^2 = \frac{36 \,.\, 2a^2}{(3a)^2\,(3a+1)} \Rightarrow a = \frac{23}{6},\ b = \frac{23}{3}$$

Thus $$p = P\{T \le 10\} = \frac{B_X(23/6,\ 23/3)}{B(23/6,\ 23/3)},\quad x = \frac{10-6}{12\ 6} = \frac{2}{3}.$$

This result may be read from Tables of Incomplete Beta function.

2*. If X is $B_{\text{II}}\,(a, b)$, then $m = (a-1)/(c-2)$; writing $\mu = a/c$ this gives $m = (\mu c - 1)/(c-2)$.

$$\therefore \quad m - \mu = \frac{\mu c - 1}{c-2} - \mu = \frac{2\mu - 1}{c-2};\quad \text{Set:}\quad f(\mu) = |m - \mu| = \frac{2|\mu - \frac{1}{2}|}{c-2} \qquad \text{...(1)}$$

The function f defined by (1) is continuous in the interval [0, 1] and hence must attain its supremum (and infimum) at $\mu = 0$ or $\mu = 1$ (and $\mu = 1/2$). Thus,

$$\text{Sup}\ |m - \mu| = 1/(c-2), \quad \text{Inf}\ |m - \mu| = 0.$$

3*. As usual, $F(t) = P(X \le t)$, so

$$G_Y(t) = P\{(1-X) \le t\} = P\{X \ge (1-t)\} = 1 - P\,(X < 1-t\} = 1 - F(1-t).$$

$$\therefore \quad g(t) = G'(t) = F'(1-t) = f(1-t).$$

Now $f(t) = t^{a-1}(1-t)^{b-1}/B(a,b) \Rightarrow g(t) = f(1-t) = t^{b-1}\,(1-t)^{a-1}/B(a,b);\ 0 < t < 1.$

4*. Let $u = xy$, $v = x$ so that $x = v$, $y = u/v$. Easily we get $dx\,dy = |J|\,du dv = v^{-1}\,du dv$. The region $0 < x < 1$ and $0 < y < 1$, in the x-y plane transforms to the region $0 < u < 1$; $u < v < 1$ in the u-v plane (Draw Figure).

Since X and Y are independent, their joint probability element is

$$dP = \frac{x^{a-1}\,(1-x)^{b-1}\,dx}{B(a,b)} \cdot \frac{y^{p-1}\,(1-y)^{q-1}\,dy}{B(p,q)} = K(1-v)^{b-1}\,u^{p-1}\,(v-u)^{q-1}\,du\,dv;$$

where $K^{-1} = B(a,b)\,.\,B(p,q)$. $p + q = a$. Integrating **out** v, we get

$$dF(u) = Ku^{p-1}\,du \int_u^1 (1-v)^{b-1}\,(v-u)^{q-1}\,dv \qquad [\text{Put } v - u = (1-u)t]$$

$$= Ku^{p-1}du\,(1-u)^{b+q-1}\int_0^1 t^{q-1}\,(1-t)^{b-1}\,dt = Ku^{p-1}\,(1-u)^{b+q-1}\,du\,.\,B(b,q)$$

This gives : $f(u) = \dfrac{u^{p-1}(1-u)^{b+q-1}}{B(p, b+q)}, \quad 0 < u < 1.$

Since $[B(b,q)/B(a,b).B(p,q)] = 1/B(p,b+q)$.

Sec. 19-33. Page 579

1*. Let $X \sim B_{\text{I}}(a, b)$, then its probability differential is

$$dF_X(x) = x^{a-1}(1-x)^{b-1}\, dx / B(a,b), \quad 0 < x < 1 \qquad ...(1)$$

Put $x = 1/(1+y)$, i.e. $y = (1-x)/x$ ...(2)

Then $|dx| = |dy|/(1+y)^2$. Further $0 < x < 1 \Rightarrow 0 < y < \infty.$ ...(3)

Substitutions from (2) and (3) into (1) yield

$$dF_Y(y) = y^{b-1}\, dy / B(a,b)\,(1+y)^{a+b}, \quad 0 < y < \infty \qquad ...(4)$$

Thus $Y \sim B_{\text{II}}(a, b)$. The reciprocal of Y i.e. $Z = 1/Y$ gives

$$dF_Z(z) = z^{a-1}\, dz / (a,b)\,(1+z)^{a+b}, \quad 0 < z < \infty \qquad ...(5)$$

Thus $Z \sim B_{\text{II}}(a, b)$. It follows that corresponding to *one* variate $X \sim B_{\text{I}}(a, b)$ there corresponds a pair of B_{II}-variates.

Converse : Let $Y \sim B_{\text{II}}(b, a)$ with probability differential (4). Put $y = (1-x)/x$ from (2) then $|dy| = |dx|/x^2$ and (4) reduces to (1). That is

$$dF_X(x) = x^{a-1}(1-x)^{b-1}\, dx / B(a,b), \quad 0 < x < 1 \qquad ...(1)$$

Thus $X \sim B_{\text{I}}(a, b)$. Further, put $x = 1 - w$, to obtain from (1)

$$dF_W(w) = w^{b-1}(1-w)^{a-1}\, dw / B(a,b), \quad 0 < w < 1 \qquad ...(6)$$

Thus $W \sim B_{\text{I}}(b, a)$. It follows that corresponding to *one* variate $X \sim B_{\text{II}}(a, b)$ corresponds a pair of B_{I} variates.

2*. Let $u = x + y$, $v = x/y$, then $x = uv/(1+v)$, $y = u/(1+v)$.

$$\frac{\partial(x,y)}{\partial(u,v)} = \begin{vmatrix} \partial x/\partial u & \partial y/\partial u \\ \partial x/\partial v & \partial y/\partial v \end{vmatrix} = \begin{vmatrix} v/(1+v) & 1/(1+v) \\ u/(1+v)^2 & -u/(1/v)^2 \end{vmatrix} = \frac{-u}{(1+v)^2}.$$

Thus, $dx\,dy = [u/(1+v)^2]\,du\,dv$; Also, the region $0 < x < \infty$, $0 < y < \infty$ in the x-y plane transforms to the region $0 < u < \infty$, $0 < v < \infty$ in the u-v plane (Draw Figure).

Since X and Y are independent, the joint p.d.f. of X and Y is

$$dP(x,y) = \frac{e^{-\lambda x}\, x^{a-1}\, dx}{\Gamma(a)/\lambda^a} \cdot \frac{e^{-\lambda y}\, y^{b-1}\, dy}{\Gamma(b)/\lambda^b} = K\lambda^{a+b}\, x^{a-1}\, y^{b-1}\, e^{-\lambda(x+y)}\, dx\, dy, \quad [K^{-1} = \Gamma(a)\,\Gamma(b)]$$

i.e. $$dP_1(u,v) = K\lambda^{a+b}\, e^{-\lambda u}\, u^{a+b-1}\, v^{a-1}\,(1+v)^{-(a+b)}\, du\, dv.$$

If $f(u, v)$ is the joint density of U, V this can be expressed as

$$f(u, v) = \left(\frac{\lambda^{a+b} e^{-\lambda u} u^{a+b-1}}{\Gamma(a+b)}\right)\left(\frac{v^{a-1}}{B(a, b)(1+v)^{a+b}}\right) = f_1(u) \cdot f_2(v), \quad 0 < u, \ v < \infty.$$

If follows that U and V are independent variates with p.d.f.'s as stated.

Note. "Quotient of two independent gamma variates is *a* B_{II}-variate" is again exemplified. (Ex.1)

3*. Let $u = x + y$, $v = x/y$, so $x = uv/(1 + v)$, $y = u/(1 + v)$. $| J | = u/(1 + v)^2$

The region $x > 0$, $y > 0$, $x + y \le 1$ transforms to $v \ge 0$, $0 \le u \le 1$.

The joint p.d.f. of U and V is $g(u, v) = f(u, v) \mid J \mid$.

$$\therefore \quad g(u, v) = \frac{24uv}{(1+v)} \frac{u}{(1+v)} \frac{u}{(1+v)^2} = (4u^3)\left(\frac{6v}{(1+v)^4}\right)$$

$$= \left[\frac{u^{4-1}.(1-u)^{1-1}}{B(4,1)}\right]\left[\frac{v^{2-1}}{B(2,2)(1+v)^{2+2}}\right]; \quad 0<u<1, \ u>0.$$

Thus U and V are independent ; $U \sim B_I(4, 1)$ and $V \sim B_{II}(2, 2)$.

Sec. 19-43. Page 582

1*. Here $\quad E(X^k) = \int_{-\infty}^{\infty} f(x)\, x^k\, dx = \int_{-\infty}^{\infty} \frac{1}{2} \lambda\, e^{-\lambda |x|}\, x^k\, dx.$

When k is odd, the integrand is odd function in $]-\infty, \infty[$ and hence $E(X^{2n-1}) = 0$. Thus $E(X) = 0$; whence simple and central moments coincide. Note $\mu = 0$ is axiomatic.

$$\therefore \quad \mu_{2k} = \lambda \int_0^{\infty} x^{2k} e^{-\lambda x} dx = \frac{\Gamma(2k+1)}{\lambda^{2k}} = \frac{2k!}{\lambda^{2k}}.$$

So, $\quad E(X^2) = \mathrm{Var}(X) = 2/\lambda^2$, $E(x^4) = 24/\lambda^4$. Further $E(Y) = a + cE(X^2) = a + (2c/\lambda^2)$

$$E(Y^2) = E(a^2 + b^2X^2 + c^2X^4 + 2abX + 2bcX^3 + 2acX^2) = a^2 + (b^2 + 2ac)(2/\lambda^2) + (24c^2/\lambda^4).$$

$$\therefore \quad \mathrm{Var}(Y) = (2b^2/\lambda^2) + (20c^2/\lambda^4).$$

$$E(XY) = E(aX + bX^2 + cX^3) = (2b/\lambda^2) \Rightarrow \sigma_{XY} = E(XY) - E(X)E(Y) = 2b/\lambda^2.$$

Thus, $\quad \rho_{XY} = \sigma_{XY'} \sigma_X \sigma_Y = \lambda b\Big/\sqrt{(b^2\lambda^2 + 10c^2)}.$

Check. $c = 0 \Rightarrow \rho(X, Y) = \pm 1$, as it should be.

2*. The density $f(x, \theta)$ is an even function of x, hence all the odd order simple moments are zero. So $E(X) = 0 = \mu$. The even-order central moments are given by

$$\mu_{2n} = \frac{\lambda^{\theta}}{2\Gamma(1+\theta)} \int_{-\infty}^{\infty} x^{2n} e^{-\lambda |x|^{1/\theta}} dx = \frac{\lambda^{\theta}}{\theta\Gamma(\theta)} \int_0^{\infty} x^{2n} \exp(-\lambda x^{1/\theta})\, dx \qquad [\text{Put } x = z^{\theta}]$$

$$= \frac{\lambda^{\theta}}{\Gamma(\theta)} \int_0^{\infty} z^{(2n+1)\theta - 1} e^{-\lambda z} dz = \frac{\lambda^{\theta}}{\Gamma(\theta)} \cdot \frac{\Gamma[(2n+1)\theta]}{\lambda^{(2n+1)\theta}} = \frac{\Gamma[(2n+1)\theta]}{\lambda^{2n\theta}\Gamma(\theta)}.$$

In particular, $\mu_2 = \Gamma(3\theta)/\lambda^{2\theta}\,\Gamma(\theta),\ \mu_4 = \Gamma(5\theta)/\lambda^{4\theta}\,\Gamma(\theta) \Rightarrow \beta_2 = \mu_4/\mu_2^2 = \Gamma(5\theta)\,\Gamma(\theta)/[\Gamma(3\theta)]^2$.

M.G.F. $$M(t:X) = \sum_{n=0}^{\infty} \frac{t^{2n}}{2n!}\mu_{2n} = \frac{1}{\Gamma(\theta)} \sum_{n=0}^{\infty} \frac{t^{2n}}{2n!} \frac{\Gamma[(2n+1)\theta]}{\lambda^{2n\theta}}.$$

Special Cases : $f(x, 1) = \frac{1}{2}\lambda e^{-\lambda|x|}$ [Lap (0, λ)]

$f(x, \frac{1}{2})$ is $N(0, 1/2\lambda)$; $f(x, 0) = \frac{1}{2}$, $|x| < 1$. (Uniform)

Sec. 19-60. Page 589

1*. Here $f(x, y) = f_1(x)\, f_2(y) = (24\sqrt{xy})^{-1} \exp[-\sqrt{x}/3 - \sqrt{y}/2]$, $x \ge 0$, $y \ge 0$.

Let $z = x/y$, $w = y$, then $|\partial(z, w)/\partial(x, y)| = y^{-1}$. Thus, the joint distribution of Z, W is

$$f(z, w) = (24\sqrt{z})^{-1} \exp\left\{-\sqrt{w}\left(\frac{1}{2} + \frac{\sqrt{z}}{3}\right)\right\}$$

$$\therefore \quad f(z) = (24\sqrt{z})^{-1} \int_0^{\infty} \exp\{-\sqrt{w}\,[(1/2) + \sqrt{z}/3]\}\, dw \qquad [\sqrt{w} = u]$$

$$= (12\sqrt{z})^{-1}\ [(1/2) + \sqrt{z}/3]^{-2},\ z > 0. \qquad \text{[gamme integral]}$$

$$p = P(Z < 1) = \int_0^1 (12\sqrt{z})^{-1}\ [(1/2) + \sqrt{z}/3]^{-2}\, dz. \qquad [\text{Put } (1/2) + \sqrt{z}/3 = t]$$

$$= \int_{1/2}^{5/6} \frac{dt}{2t^2} = \frac{2}{5}.$$

2*. We firstly obtain the "Survival c.d.f. of Weibull (k; μ, a).

$$P(X > t) = \int_t^{\infty} f(x)\, dx = \int_{t_0}^{\infty} e^{-t}\, du \qquad \left[u = \left(\frac{x-\mu}{a}\right)^k,\ t_0 = \left(\frac{t-\mu}{a}\right)^k\right]$$

So, $$P(X > t) = \exp[-t_0] \qquad \text{...(2)}$$

$$\therefore \quad P\{Y > y\} = P\{X_1 > y, X_2 > y, \ldots, X_n > y\}$$
$$= P(X_1 > y)\,.\,P(X_2 > y) \ldots P(X_n > y) = [P(X > y)]^n$$
$$= \exp[-ny_0] = \exp\{-[n^{1/k}(y-\mu)/a]^k\} \text{ by (2),} \qquad \text{...(3)}$$

Result (3) implies that $Y \sim$ Weibull $[k\ ;\ \mu,\ a/(n)^{1/k}]$.

3*. Firstly, we determine the p.d.f. of Z. Let G be the c.d.f. of Z; then

$G(z) = P(Z \le z) = P\{\max.\ (X_1, X_2, \ldots, X_n) \le z\} = P\{X_i \le z,\ i = 1, 2, \ldots, n\}$

$= P(X_1 \le z)\, P(X_2 \le z) \ldots P(X_n \le z) = [F(z)]^n$ [∵ X_i are i.i.d]

where $F(z)$ is the common c.d.f. of X's. Since $X \sim$ Expo (λ), $F(x) = 1 - e^{-\lambda x}$, $x \ge 0$, and so

$$\therefore \quad G(z) = [1 - e^{-\lambda z}]^n,\ g(z) = G'(z) = n\lambda e^{-\lambda z}(1 - e^{-\lambda z})^{n-1},\ z \ge 0.$$

$$\mathrm{E}(e^{tZ}) = \int_0^{\infty} e^{tz}\,.\,n\lambda\, e^{-\lambda z}(1 - e^{-\lambda z})^{n-1}\, dz = n\int_0^1 u^{n-1}(1-u)^{-t/\lambda}\, du. \quad [u = 1 - e^{-\lambda z}]$$

$$\therefore \quad M(t : Z) = nB[n,\ 1 - (t/\lambda)] = n\Gamma(n)\, \Gamma[1 - (t/\lambda)]/\Gamma[n + 1 - (t/\lambda)].$$

Let $k = 1 - (t/\lambda)$; use $\Gamma(n + k) = (n + k - 1)\,(n + k - 2) \ldots (k - 1)\, k\,.\,\Gamma(k)$; whence

$$[M(t:Z)]^{-1} = \frac{(n+k-1)(n+k-2)\dots(k+1)k}{n(n-1)(n-2)\dots2.1} = \left(1+\frac{k-1}{n}\right)\left(1+\frac{k-1}{n-1}\right)\left(1+\frac{k-1}{n-2}\right)\dots\left(1+\frac{k-1}{2}\right)\cdot\frac{k}{1}$$

$$\therefore \quad M(t:Z) = \left[\left(1-\frac{\theta}{n}\right)\left(1-\frac{\theta}{n-1}\right)\dots\left(1-\frac{\theta}{2}\right)(1-\theta)\right]^{-1} \quad \dots(1)$$

where $t/\lambda = \theta = 1 - k$. We now determine the m.g.f. of Y.

$$M(t:Y) = \mathsf{E}\{e^{t[X_1+(1/2)X_2+\dots+(1/n)X_n]}\} = \mathsf{E}(e^{tX_1})\,\mathsf{E}(e^{(1/2)tX_2})\dots\mathsf{E}[e^{(t/n)X_n}] \qquad (X_i \text{ are i.i.d.})$$

$$= \left(1-\frac{t}{\lambda}\right)^{-1}\left(1-\frac{t}{2\lambda}\right)^{-1}\dots\left(1-\frac{t}{n\lambda}\right)^{-1} = \left[(1-\theta)\left(1-\frac{\theta}{2}\right)\dots\left(1-\frac{\theta}{n}\right)\right]^{-1} \quad \dots(2)$$

From (1) and (2), it follows that $M(t:Y) = M(t:Z)$. Since m.g.f. uniquely determines the c.d.f. it follows that Y and Z have the same distribution.

4*. By definition $\mu_r = \mathsf{E}(X^r) = \int_{-\infty}^{\infty} x^r f(x)\,dx.$ [Assuming $\mathsf{E}(X) = 0$] ...(2)

Multiply stated Eq. (1) by x^{2n+1} and integrate over $(-\infty, \infty)$ to get

$$\int_{-\infty}^{\infty} x^{2n+2} f(x)\,dx = \int_{-\infty}^{\infty} x^{2n+1}(b_0 + b_2x^2 + b_4x^4) f'(x)\,dx$$

or $$\mu_{2n+2} = [x^{2n+1}(b_0 + b_2x^2 + b_4x^4) f(x)]_{-\infty}^{\infty} - \int_{-\infty}^{\infty}[(2n+1)b_0x^{2n} + (2n+3)b_2x^{2n+2} + (2n+5)b_4x^{2n+4}] f(x)\,dx.$$

Since $f(-\infty) = 0 = f(\infty)$, using (2), this equation reduces to

$$b_0(2n+1)\mu_{2n} + b_2(2n+3)\mu_{2n+2} + b_4(2n+5)\mu_{2n+4} = -\mu_{2n+2}.$$

5*. $$\phi^{(r)} = \frac{d^r\phi}{d\theta^r} = \frac{d^r\mathsf{E}(e^{\theta X})}{d\theta^r} = \mathsf{E}(X^r e^{\theta X}) = \int_{-\infty}^{\infty} x^r e^{\theta x} f(x)\,dx \qquad [\theta = it] \quad \dots(i)$$

Put $df/dx = f'$, multiply the Pearson differential equation by e^{θ} and integrate over $(-\infty, \infty)$ to obtain

$$\int_{-\infty}^{\infty} e^{\theta x}(b_0 + b_1x + b_2x^2) f'(x)\,dx = \int_{-\infty}^{\infty} e^{\theta x}(a + x) f(x)\,dx.$$

Assume that integrals vanish at either terminal limit, use (i) for R.H.S. to get

$$[e^{\theta x}(b_0 + b_1x + b_2x^2) f(x)]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} e^{\theta x} f(x)[\theta(b_0 + b_1x + b_2x^2) + (b_1 + 2b_2x)]\,dx = a\phi + \phi'$$

or $$-[\theta(b_0\phi + b_1\phi' + b_2\phi'') + (b_1\phi + 2b_2\phi')] = a\phi + \phi', \qquad \text{[by (i)]}$$

i.e. $$b_2\theta\phi'' + (1 + 2b_2 + b_1\theta)\phi' + (a + b_1 + b_2\theta)\phi = 0. \quad \dots(1)$$

Differentiating by Leibnitz Rule n time w.r.t. 'θ' we get

$$b_2[\theta\phi^{n+2} + n\phi^{n+1}] + [\phi^{n+1}(1 + 2b_2 + b_1\theta) + nb_1\phi^n] + [\phi^n(a + b_1 + b_0\theta) + nb_0\phi^{n-1}] = 0.$$

Now put $\theta = 0$ so that $\phi^r(0) = \mu'_r$. And if we shift the origin to the mean, then $\phi^r(0) = \mu'_r = \mu_r$. Collecting the like terms the last Eq. gives the moment recurrence formula.

$$[(n+2)b_2+1]\,\mu_{n+1}+[(n+1)b_1+a]\mu_n+nb_0\mu_{n-1}=0. \quad \ldots(2)$$

Cumulant generating function. By definition, $\psi=\ell n\phi$ or $\phi=e^{\psi}$, so that

$$\phi'=e^{\psi}\,\psi',\ \phi''=e^{\psi}\,(\psi''+\psi'^2)$$

Substituting these values into (1) and cancelling $e^{\psi}(\neq 0)$, we get

$$b_2\theta\,(\psi''+\psi'^2)+(1+2b_2+b_1\theta)\psi'+(a+b_1+b_0\theta)=0. \quad \ldots(3)$$

To obtain cumulants recurrence formula we differentiate (3) r times and use Leibnitz Rule :

$$b_2\theta\,(D_r\psi''+D_r\psi'^2)+\binom{r}{1}b_2\,(D_{r-1}\psi''+D_{r-1}\psi'^2)+(1+2b_2+b_1\theta)\,\psi^{r+1}+\binom{r}{1}b_1\psi^r=0,$$

where $D_r=d^r/d\theta^r$. Now put $\theta=0$, use $(k_r=D_r\psi)_{\theta=0}$; this gives

$$[1+(r+2)b_2]k_{r+1}+rb_1k_r+rb_2\,(D_{r-1}\psi'^2)_{\theta=0}=0. \quad \ldots(4)$$

To evaluate the last term, write $\psi'^2=\psi_1\psi_1$, $(\psi_1=d\psi/d\theta)$ and apply Leibnitz Rule :

$$X_{r-1}(\psi_1\psi_1)=(\psi_1)_{r-1}\psi_1+\binom{r-1}{1}(\psi_1)_{r-2}(\psi_1)_1+\ldots+\binom{r-1}{j}(\psi_1)_{r-1-j}(\psi_1)_j+\ldots+\binom{r-1}{r-1}(\psi_1)_0(\psi_1)_{r-1}.$$

Putting $\theta=0$, noting $(\psi_1)_0=k_1=0$, as $\mu'_1=0$, we get

$$(D_{r-1}\psi_1^{-2})_0=\binom{r-1}{1}k_2k_{r-1}+\ldots+\binom{r-1}{j}k_{j+1}+k_{r-j}+\ldots+\binom{r-1}{r-2}k_2k_{r-1}$$

Substituting this value in (4), we get the result as stated.

Chapter 20 : Bivariate Normal Distribution

Sec. 20-60. Page 609

1*. $$J^{-1}=\begin{bmatrix}\partial u/\partial x & \partial v/\partial x\\ \partial u/\partial y & \partial v/\partial y\end{bmatrix}=\begin{bmatrix}\sigma_1^{-1} & 0\\ -\rho/\sigma_2 & \sqrt{1-\rho^2}/\sigma_2\end{bmatrix}$$

$$\therefore\quad J=\sigma_1\sigma_2/\sqrt{1-\rho^2},\qquad u^2+v^2=(x^2/\sigma_1^2)-2xy\,\rho/\sigma_1\sigma_2+(y^2/\sigma_2^2).$$

$$\therefore\quad f(u,v)=\frac{e^{-(u^2+v^2)/2(1-\rho^2)}}{2\pi(\sqrt{1-\rho^2})^2}=f_1(u).f_2(v)\Rightarrow U\sim N(0,1-\rho^2)\text{ and }V\sim N(0,1-\rho^2)\text{ are indep.}$$

Note. M.G.F. method is very convenient. Try it.

2*. If $(X, Y)\sim$ BVN $(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2; \rho)$, then $f(x,y)=(2\pi\sigma_1\sigma_2\sqrt{1-\rho^2})^{-1}e^{-Q/2}$

where $Q=\{(x-\mu_1)^2/\sigma_1^2-2\rho(x-\mu_1)(y-\mu_2)/\sigma_1\sigma_2+(y-\mu_2)^2/\sigma_2^2\}/(1-\rho^2)$.

Comparison provides : $\mu_1=2$, $\mu_2=-3$, $C=(2\pi\sigma_1\sigma_2\sqrt{1-\rho^2})^{-1}$

$$\frac{16}{216}=\frac{1}{2(1-\rho^2)\,\sigma_1^2},\ \frac{9}{216}=\frac{1}{2(1-\rho^2)\sigma_2^2},\ \frac{2\rho}{2\sigma_1\sigma_2\,(1-\rho^2)}=\frac{12}{216}$$

$\therefore \quad (1-\rho^2)\sigma_1^2 = 27/4,\ (1-\rho^2)\sigma_2^2,\ \sigma_1\sigma_2(1-\rho^2) = 18\rho.$

So $\quad (1-\rho^2)^2\sigma_1^2\sigma_2^2 = 81 = (18\rho)^2 \Rightarrow \rho^2 = 1/4.$

$\therefore \quad \sigma_1 = 3,\ \sigma_2 = 4, C = [2\pi.3.4(\sqrt{3}/2)]^{-1} = 1/12\sqrt{3}\pi.$

Thus, the given BVN is $N(2, -3, 16;\ 1/2)$.

3*. Obviously, $\mu_X = 0 = \mu_Y$. Hence we compare (1) with

$$f = (2\pi\sigma_1\sigma_2\sqrt{1-\rho^2})^{-1}\exp\{[(x^2/\sigma_1^2)+(y^2/\sigma_2^2)-(2\rho xy/\sigma_2\sigma_2)]/2(1-\rho^2)\}$$

just to find $\sigma_1\sigma_2\sqrt{1-\rho^2} = \sqrt{3},\ \sigma_1^2(1-\rho^2) = 3/4, \sigma_2^2(1-\rho^2) = 3; \rho/\sigma_1\sigma_2(1-\rho^2) = \frac{1}{2}.$

Solving : $\sigma_2^2 = 4,\ \sigma_1^2 = 1,\ \rho^2 = \frac{1}{4}$. Thus Corr $(X, Y) = \rho = \pm\frac{1}{2}$.

4*. $P(XY > 0) = P\{(X > 0, Y > 0) \uplus (X < 0, Y < 0)\} = P(X > 0, Y > 0) + P(X < 0, Y < 0)\}$
$= 2P(X > 0, Y > 0)$. (by symmetry) [use Exp. 20-3]
$= 2[(1/4) + (1/2\pi)\sin^{-1}\rho] = (1/2) + \pi^{-1}\sin^{-1}\rho.$

The other part follows trivially by noting $\sin^{-1}\theta + \cos^{-1}\theta = \pi/2$.

5*. Recall : $\quad (Y \mid x) \sim N(m_2,\ \sigma_2'^2)$, where

$$m_2 = \mu_2 + \rho(\sigma_2/\sigma_1)(x-\mu_2) = 10 + \rho(5/1)(5-5) = 10, \sigma_2'^2 = \sigma_2^2(1-\rho^2) = 25(1-\rho^2).$$

Put $Y^* = (Y - m_2)/\sigma_2'$, i.e. $Y = 10 + \sigma_2' Y^*$, so that, by hypothesis (writing $\sigma = \sigma'^2$)

$$0.954 = P\{4 < 10 + \sigma Y^* < 16\} = P\{-6/\sigma < Y^* < 6/\sigma\} = 2\psi(6/\sigma)$$

Thus $\quad \psi(6/\sigma) = 0.4770 \Rightarrow 6/\sigma = 2$ or $\sigma = 3$ [From Tables]

As such, $\quad 25(1-\rho^2) = 9 \Rightarrow \rho = 4/5$ (since $\rho > 0$).

(ii) If $\rho = 0$, X and Y being jointly normal, are certainly independent. Thus $Z = (X + Y) \sim N(15, 26)$. Put $Z^* = (Z - 15)/\sqrt{26}$ i.e. $Z = 15 + \sqrt{26}\,Z^*$. Now

$$p = P\{X + Y \le 16\} = P\{Z \le 16\} = P\{15 + \sqrt{26}\,Z^* \le 16\} = P\{Z^* \le 1/\sqrt{26}\}$$
$$= 0.5000 + \psi(0.1961161) = 0.50 + 0.0753 = 0.5753.$$

6*. $E(Y \mid X = 7) = \mu_2 + \rho(\sigma_2/\sigma_1)(x - \mu_1) = 1 + (3/4)(7 - 3) = 4.$

Var $(Y \mid X = 7) = \sigma_2^2(1-\rho^2) = 25(1 - 9/25) = 16$. Thus $Y/x \sim N(4, 16)$.

Put $\quad Z = (Y - 4)/4$, i.e. $Y = 4 + 4Z$ where $Z \sim N(0, 1)$. So

$P\{3 < Y < 8 \mid x = 7\} = P\{3 < (4 + 4Z) < 8\} = P\{-1/4 < Z < 1\}$
$= \psi(0.25) + \psi(1) = 0.0987 + 0.3413 = 0.44,$ [Tables of $N(0, 1)$ Area]

(b) $\quad E(X \mid y) = \mu_1 + \rho(\sigma_1/\sigma_2)(y - \mu_2) = 3 + (12/25)(-4 - 1) = 3/5.$

Var $(X \mid y) = \sigma_1^2(1-\rho^2) = 16[1 - (9/25)] = (16/5)2.$

Thus $(X \mid y)\ N[(3/5), (16/5)^2]$. Now put $Z = (X - 3/5)/(16/5)$, i.e. $X = (3/5) + (16Z/5)$.

$P(-3 < X < 3 \mid Y = -4) = P\{-3 < [(3/5) + (16Z/5)] < 3\} = P\{-1.125 < Z < 0.75\}$
$= \psi(1.125) + \psi(0.75) = 0.3708 + 0.2734 = 0.6442.$ [$N(0,1)$-tables]

7*. Compare the given p.d.f. with that of BVN $(\mu_1, \mu_2, \sigma_1^2; \sigma_1^2; \sigma)$:

$$f(x, y) = (2\pi\sigma_1\sigma_2\sqrt{1-\rho^2})^{-1} \exp\left\{-\left[\left(\frac{x-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{x-\mu_1}{\sigma_1}\right)\left(\frac{y-\mu_2}{\sigma_2}\right) + \left(\frac{y-\mu_2}{\sigma_2}\right)^2\right]\Big/2(1-\rho^2)\right\}$$

Coeff. of exponential : $\lambda / 2\pi\,\sigma_1\sigma_2 = 1/6\,\pi\sqrt{3},\ \lambda = (1-\rho^2)^{-1/2}$.

Coeff. of x^2 : $\frac{1}{2}(\lambda^2/\sigma_1^2) = 2/27$, Coeff. of $y^2 : \frac{1}{2}(\lambda^2/\sigma_2^2) = \frac{1}{6}$; Coeff. of $xy : \rho\lambda^2/\sigma_1\sigma_2 = -1/9$,

Coeff. of x : $\lambda^2\left[\frac{\mu_1}{\sigma_1^2} - \rho\frac{\mu_2}{\sigma_1\sigma_2}\right] = \frac{5}{9}$; Coeff. of $y : \lambda^2\left[\frac{\mu_2}{\sigma_2^2} - \rho\frac{\mu_1}{\sigma_1\sigma_2}\right] = \frac{2}{3}$

Constant : $\frac{1}{2}\lambda^2\left[\frac{\mu_1^2}{\sigma_1^2} + \frac{\mu_2^2}{\sigma_2^2} - \frac{2\rho\mu_1\mu_2}{\sigma_1\sigma_2}\right] = \frac{7}{6}$

From first three relations : $\sigma_1 = 3,\ \sigma_1 = 2,\ \rho = -\frac{1}{2},\ (\lambda = 2/\sqrt{3})$

From the 5th and 6th relations : $\mu_1 = 3,\ \mu_2 = 1$.

These values satisfy the 4th and 7th relations. Now regression of the mean of Y on X is

$$\mathrm{E}(Y \mid X) = \mu_2 + \rho(\sigma_2/\sigma_1)(x - \mu_1) = 2 - (x/2).$$

Since this is a linear estimate, the associated M.S. error is given by

$$\varepsilon = \sigma_2^2(1-\rho^2) = 4[1 - (\tfrac{1}{4})] = 3.$$

8*. We utilize the m.g.f. of BVN. Let $e^X = U,\ e^Y = V$; then

$$\mathrm{E}(U^{t_1}V^{t_2}) = \mathrm{E}(e^{t_1X + t_2Y}) = \exp\{\mu_1 t + \mu_2 t + \tfrac{1}{2}(\sigma_1^2 t_1^2 + 2\rho\sigma_1\sigma_2 t_1 t_2 + \sigma_2^2 t_2^2)\}$$

$$\therefore \quad \mathrm{E}(U) = e^{\mu_1 + \frac{1}{2}\sigma_1^2},\ \mathrm{E}(U^2) = e^{2(\mu_1 + \sigma_1^2)};\ \mathrm{E}(VU) = e^{\mu_1 + \mu_2}\exp[\tfrac{1}{2}(\sigma_1^2 + 2\rho\sigma_1\sigma_2 + \sigma_2^2)].$$

Thus, $\mathrm{Var}(U) = \mathrm{E}(U^2) - \mathrm{E}^2(U) = e^{2\mu_1 + \sigma_1^2}(e^{\sigma_1^2} - 1)$. $\mathrm{Var}(V) = e^{2\mu_2 + \sigma_2^2}(e^{\sigma_2^2} - 1)$.

$$\mathrm{Cov}(U, V) = \mathrm{E}(UV) - \mathrm{E}(U)\mathrm{E}(V) = e^{\mu_1 + \mu_2}\, e^{(1/2)(\sigma_1^2 + \sigma_2^2)}\,(e^{\rho\sigma_1\sigma_2} - 1).$$

Thus, $$\mathrm{Corr}(U, V) = \frac{\sigma_{UV}}{\sigma_U\sigma_V} = \frac{(e^{\rho\sigma_1\sigma_2} - 1)}{[(e^{\sigma_1^2} - 1)(e^{\sigma_2^2} - 1)]^{1/2}}.$$

9*. (i) $$M(t_1, t_2) = \mathrm{E}(e^{t_1U + t_2V}) = \mathrm{E}\{e^{t_1(X + \lambda Z)}\, e^{t_2(Y + \lambda Z)}\} = \mathrm{E}(e^{t_1X})\,\mathrm{E}(e^{t_2Y})\,\mathrm{E}[e^{\lambda(t_1 + t_2)Z}]$$

$$= e^{t_1^2/2}.e^{t_2^2/2}.e^{\lambda^2(t_1 + t_2)^2/2} = \exp[\tfrac{1}{2}(1+\lambda^2)t_1^2 + \tfrac{1}{2}(1+\lambda^2)t_2^2 + \lambda^2 t_1 t_2] \quad \ldots(1)$$

For BVN $(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2; \rho)$, the m.g.f. is

$$M(t_1, t_2) = \exp[\mu_1 t_1 + \mu_2 t_2 + \tfrac{1}{2}(\sigma_1^2 t_1^2 + 2\rho\sigma_1\sigma_2 t_1 t_2 + \sigma_2^2 t_2^2)] \quad \ldots(2)$$

Comparing (1) and (2) : $(U, V) \sim \mathrm{BVN}(0, 0, 1 + \lambda^2, 1 + \lambda^2; \rho)$

where $\rho = \lambda^2/(1+\lambda^2)$, $(\lambda^2 = \rho\sigma_1\sigma_2)$. Obviously, when $\rho = 1/2$, $\lambda = \pm 1$.

(ii) Let $U_1 = \mu_1 + (\sigma_1 U/\sqrt{1+\lambda^2})$, $\quad V_1 = \mu_2 + (\sigma_2 V/\sqrt{1+\lambda^2})$. ...(3)

Then $\mathrm{E}(U_1) = \mu_1$, $\mathrm{Var}\,(U_1) = \sigma_1^2$, $\mathrm{E}(V_1) = \mu_2$, $\mathrm{Var}\,(V_1) = \sigma_2^2$

$$\mathrm{Cov}\,(U_1, V_1) = \frac{\sigma_1\sigma_2}{1+\lambda^2}\mathrm{Cov}\,(U, V) = \frac{\sigma_1\sigma_2}{1+\lambda^2}\lambda^2 = \lambda^2, \quad \mathrm{Corr}\,(U_1, V_1) = \frac{\lambda^2}{1+\lambda^2}.$$

Thus, mapping (3) gives the required result.

10*. Given distribution is the mixture of BVN and trinomial law. Now

$$M(t_1, t_2) = \mathrm{E}(e^{t_1X+t_2Y}) = k\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{t_1X+t_2Y} f(x, y) + (1-k)\sum_{x=0}^{n}\sum_{y=0}^{n-x} e^{t_1X+t_2Y}\; p(x, y).$$

$$= k\,[\text{m.g.f. of BVN}] + (1-k)\,[\text{m.g.f. of Trinomial law}].$$

$$= k\exp\left[t_1\mu_1 + t_2\mu_2 + \tfrac{1}{2}(t_1^2\sigma_1^2 + 2\rho t_1t_2\sigma_1\sigma_2 + t_2^2\sigma_2^2)\right] + (1-k)\,(1-p-q+pe^{t_1}+qe^{t_2})^n.$$

$$\therefore \quad M(t_1, 0) = k\exp\left[t_1\mu_1 + \tfrac{1}{2}\sigma_1^2 t_1^2\right] + (1-k)\,(1-p+pe^t)^n.$$

Thus, X also has the mixed density consisting of $N(\mu_1, \sigma_1^2)$ and bin (n, p).

$$f(x) = k(\sqrt{2\pi}\sigma_1)^{-1}\exp[-(x-\mu_1)^2/2\sigma_1^2];\ -\infty < x < \infty, \text{ omitting integers}$$

$$= (1-k)\,{}^nC_x\, p^x(1-p)^x;\ x = 0, 1, 2, \ldots$$

$$\therefore \quad \mathrm{E}(X) = k\mu_1 + (1-k)\,np;\quad \mathrm{E}(X^2) = k(\sigma_1^2+\mu_1^2) + (1-k)\,[np + n(n-1)p^2].$$

$$\therefore \quad \mathrm{Var}\,(X) = k(\sigma_1^2+\mu_1^2) + (1-k)\,[np + n(n-1)p^2] - [k\mu_1 + (1-k)np]^2$$

$$\mathrm{E}(XY) = \partial^2 M(0,0)/\partial t_1\partial t_2 = k[\rho\sigma_1\sigma_2 + \mu_1\mu_2] + (1-k)\,n(n-1)\,pq$$

$$\therefore \quad \mathrm{Cov}\,(X, Y) = \mathrm{E}(XY) - \mathrm{E}(X)\,\mathrm{E}(Y)$$

$$= k[\rho\sigma_1\,\sigma_2 + \mu_1\mu_2] + (1-k)n\,(n-1)\,pq - [k\mu_1 + (1-k)np]\,[k\mu_2 + (1-k)nq].$$

Now for BVN, $(X \mid y) \sim N[\mu_1 + \rho(\sigma_1/\sigma_2)\,(y-\mu_2),\ \sigma_1^2\,(1-\rho^2)]$

and for Tri-nomial, $(X \mid y) \sim$ bin $[n-y,\ p/(1-q)]$.

Hence, taking into account the factors k and $(1-k)$, we get

$$\mathrm{E}(X \mid y) = k[\mu_1 + \rho\sigma_1\sigma_2^{-1}\,(y-\mu_2)] + (1-k)\,(n-y)\,p(1-q)^{-1}$$

$$\mathrm{E}(X^2 \mid y) = k\{\sigma_1^2\,(1-\rho^2) + [\mu_1 + \rho\sigma_1\sigma_2^{-1}\,(y-\mu_2)]^2\}$$

$$+ (1-k)\,\{(n-y)\,p(1-p-q)\,(1-q)^{-2} + (n-y)^2\,p^2\,(1-q)^{-2}\}. \quad [\because\ \mu_2' = \mu_2 + \mu_1'^2]$$

$$\therefore \quad \mathrm{Var}\,(X \mid y) = \mathrm{E}(X^2 \mid y) - \mathrm{E}^2\,(X \mid y).$$

We do not attempt to simplify this expression.

11*. $\int_{-\infty}^{\infty} f(x, y)\, dy = (2\pi)^{-1} e^{-x^2/2} \int_{-\infty}^{\infty} e^{-y^2/2} . [1 + xye^{-y^2/2} e^{(1-x^2/2)}]\, dy$

[2nd term is odd function of y]

$$= \frac{e^{-\frac{1}{2}x^2}}{\sqrt{2\pi}} \left\{ \int_{-\infty}^{\infty} \frac{\exp^{(-y^2/2)}}{\sqrt{2\pi}} dy + 0 \right\} = \frac{e^{-x^2/2}}{\sqrt{2\pi}}, \quad -\infty < x < \infty \qquad \ldots(1)$$

$$\therefore \quad \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x, y)\, dy\, dx = \int_{-\infty}^{\infty} \frac{e^{-x^2/2}}{\sqrt{2\pi}} dx = 1 \qquad \ldots(2)$$

Since $f(x, y) \geq 0$ and (2) yields $\sum f(x, y) = 1$, it follows that f is a density function. Incidently, (1) yields that $X \sim N(0, 1)$. Since $f(x, y)$ is symmetrical, we infer that Y is also $N(0, 1)$. Notice that X and Y are not independent, since $f(x, y) \neq f_1(x) f_2(y)$.

Remark. Each marginal density is Normal, but the joint density is not BVN.

N.B. $\int f(x, y)\, dy = f_1(x)$; $\iint f(x, y)\, dy\, dx = \int f_1(x)\, dx = 1$.

12*. If $aX + bY$ is the linear combination sought, then

$$1 = \text{Var}\,(aX + bY) = a^2\sigma_1^2 + b^2\sigma_1^2 + 2ab\,\sigma_1\sigma_2\rho = a^2 + b^2 + 2ab\,\rho \qquad \ldots(1)$$

$$\text{Cov}\,(X, aX + bY) = a\,\text{Var}\,(X) + b\,\text{Cov}\,(X, Y) = a + b\rho, \quad (\sigma_1 = \sigma_2 = 1) \qquad \ldots(2)$$

Now X and $aX + bY$ shall be independent iff $\text{Cov}\,(X, aX + bY) = 0$, since these variables are BVN. Then by (2), $a = -b\rho$ and hence by (1)

$$1 = b^2 (1 + \rho^2 - 2\rho^2) = b^2(1 - \rho^2) \Rightarrow b = \pm(1-\rho^2)^{-1/2}, \; a = \pm\rho(1-\rho^2)^{-1/2}$$

Thus, a linear combination L is possible in two ways :

$$L = (\rho X - Y)(\sqrt{1-\rho^2})^{-1} \qquad \text{or} \qquad L = (Y - \rho X)(\sqrt{1-\rho^2})^{-1}.$$

13*. The m.g.f. of $(X, Y) \sim \text{BVN}\,(0, 0, \sigma_1^2, \sigma_2^2; \rho)$, when expanded in series, gives

$$M(t_1, t_2) = \mathsf{E}(e^{t_1X + t_2Y}) = \exp\,[\tfrac{1}{2}(\sigma_1^2 t_1^2 + 2\rho\sigma_1\sigma_2 t_1 t_2 + t_2^2\sigma_2^2)]$$

$$= 1 + \tfrac{1}{2}(\sigma_1^2 t_1^2 + 2\rho\sigma_1\sigma_2 t_1 t_2 + \sigma_2^2 t_2^2 + \tfrac{1}{8}(\sigma_1^2 t_1^2 + 2\rho\sigma_1\sigma_2 t_1 t_2 + \sigma_2^2 t_2^2)^2 \qquad \ldots(1)$$

Since $\mathsf{E}(X^r . Y^s) = \text{Coeff. of } (t_1^r t_2^s / r!\,s!) \text{ in } M(t_1, t_2)$

$$\therefore \quad \mathsf{E}(X) = 0 = \mathsf{E}(Y),\; \mathsf{E}(X^2) = \sigma_1^2,\; \mathsf{E}(Y^2) = \sigma_2^2,\; \mathsf{E}(X^4) = 3\sigma_1^4, \mathsf{E}(Y^4) = 3\sigma_2^2,$$

$$\mathsf{E}(X^2Y^2) = \sigma_1^2\sigma_2^2\,(1 + 2\rho^2),\; \mathsf{E}(XY) = \rho\sigma_1\sigma_2$$

So $\quad \text{Cov}\,(X^2, Y^2) = \mathsf{E}(X^2 . Y^2) - \mathsf{E}(X^2)\,\mathsf{E}\,(Y^2) = \sigma_1^2\,\sigma_2^2\,(1 + 2\rho^2) - \sigma_1^2\,\sigma_2^2 = 2\rho^2\sigma_1^2\,\sigma_2^2.$

$$\text{Var}\,(X^2) = \mathsf{E}(X^4) - \mathsf{E}^2(X^2) = 3\sigma_1^4 - \sigma_1^4 = 2\sigma_1^4;\; \text{Var}\,(Y^2) = 2\sigma_2^4.$$

$$\therefore \quad \text{Corr}\,(X^2, Y^2) = \text{Cov}\,(X^2, Y^2) / \sqrt{\text{Var}\,(X^2) . \text{Var}\,(Y^2)} = 2\rho^2\sigma_1^2\sigma_2^2 / 2\sigma_2^2 . \sigma_1^2 = \rho^2.$$

14*. (a) Recall that m.g.f. of X and Y is

$$M(T_1, T_2) = \mathsf{E}(e^{T_1X + T_2Y}) = \exp\,[\tfrac{1}{2}(T_1^2 + 2\rho T_1 T_2 + T_2^2)]. \qquad \ldots(1)$$

Now the m.g.f. of $X + Y$ and $X - Y$ is

$$M(t_1, t_2) = \mathrm{E}[e^{t_1(X+Y)+t_2(X-Y)}] = e^{[(t_1+t_2)X+(t_1-t_2)Y]}$$

$$= \exp\{\tfrac{1}{2}[(t_1+t_2)^2 + 2\rho(t_1-t_2)(t_1+t_2) + (t_1-t_1)^2]\}, \qquad \text{[by (1)]}$$

$$= \exp\{(1+\rho)t_1^2 + (1-\rho)t_2^2\} = \exp[(1+\rho)t_1^2]\exp[(1-\rho)t_2^2] = M(t_1, 0)\, M(0, t_2)\,.$$

This shows that $U = X + Y$ and $V = X - Y$ are independently distributed with densities.

$$f_1(u) = \frac{e^{-u^2/4(1+\rho)}}{2\sqrt{\pi(1+\rho)}}, \qquad f_2(v) = \frac{e^{-v^2/4(1-\rho)}}{2\sqrt{\pi(1-\rho)}}$$

The result is otherwise obvious :

$$\mathrm{Cov}\,(X+Y, X-Y) \equiv \sigma_X^2 - \sigma_Y^2 = 1 - 1 = 0 \Rightarrow \rho_{XY} = 0 \Rightarrow \mathrm{Ind}\,(X, Y).$$

(b) When $\rho \neq 0$, we can never achieve the equality

$$M(T_1, T_2) = M(T_1, 0)\, M(0, T_2) \qquad \ldots(2)$$

so that X and Y are generally dependent. However, when $\rho = 0$.
Eq. (2) holds [put $\rho = 0$ in (1)] which shows that X and Y are independent. Conversely, independence of X and Y always leads to $\rho = 0$.

(c) $$M(t : Q) = \mathrm{E}(e^{tQ}) = (2\pi\sqrt{1-\rho^2})^{-1} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-Q/2}\,.\,e^{tQ} dx\, dy \qquad \ldots(3)$$

Let $x = u/\sqrt{1-2t},\ y = v/\sqrt{1-2t}$, so that $dx\, dy = du\, dv/(1 - 2t)$

$$\exp\left[-\frac{(1-2t)Q}{2}\right] = \exp\left[-\frac{(1-2t)}{2}\cdot\frac{x^2 - 2\rho xy + y^2}{\sqrt{1-\rho^2}}\right] = \exp\left\{-\frac{(u^2 - 2\rho uv + v^2)}{\sqrt{1-\rho^2}}\right\}$$

$$\therefore \qquad M(t : Q) = (1-2t)^{-1} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(u, v)\, du\, dv \qquad \ldots(4)$$

where $$f(u, v) = (2\pi\sqrt{1-\rho^2})^{-1} \exp[-(u^2 - 2\rho uv + v^2)/2(1-\rho^2)]$$

Since $f(u, v)$ is standard BVN-density, Eq. (4) reduces to

$M(t : Q) = (1-2t)^{-1}$ = [m.g.f. of $\chi^2_{(2)}$]. It follows that $Q \sim \chi^2_{(2)}$.

15*. $M(T_1, T_2) = \mathrm{E}(e^{T_1X+T_2Y}) = \exp[\tfrac{1}{2}(T_1^2\sigma_1^2 + 2\rho\sigma_1\sigma_2T_1T_2 + T_2^2\sigma_2^2)]$ [§20-21] $\qquad \ldots(1)$

$$m(t_1, t_2) = \mathrm{E}\{e^{t_1U+t_2V}\} = \mathrm{E}\{e^{t_1[(X/\sigma_1)+(Y/\sigma_2)]+t_2[(X/\sigma_1)-(Y/\sigma_2)]}\} = \mathrm{E}\{e^{X(t_1+t_2)\sigma_1^{-1}+Y(t_1-t_2)\sigma_2^{-1}}\}$$

$$= \exp\{\tfrac{1}{2}[(t_1+t_2)^2 + 2\rho(t_1+t_2)(t_1-t_2) + (t_1-t_2)^2]\}, \qquad \text{[by (1)]}$$

$$= [\exp(1+\rho)t_1^2]\ [\exp(1-\rho)t_2^2] = m(t_1, 0)\,.\,m(0, t_2)$$

We thus see that U and V are independent distributed, and

$$U \sim N[0, 2(1+\rho)] \text{ and } V \sim N[0, 2(1-\rho)].$$

Trivially : $\mathrm{Cov}\,(U, V) = (\sigma_1^2/\sigma_1^2) - (\sigma_2^2/\sigma_2^2) = 0$ and conclusion flows readily.

16*. Since $X \sim N(0, \sigma_1^2)$, $Y \sim N(0, \sigma_2^2)$, so

$$\mu_1 = \mu_2 = 0, \quad \mathrm{E}(X^2) = \sigma_1^2, \mathrm{E}(Y^2) = \sigma_2^2, \quad \mathrm{E}(X^4) = 3\sigma_1^4, \mathrm{E}(Y^4) = 3\sigma_2^4$$

The conditional and marginal densities give

$$\mathrm{E}(X \mid y) = (\rho\sigma_1/\sigma_2)y; \ \mathrm{Var}\,(X \mid y) = \sigma_1^2\,(1-\rho^2); \qquad \ldots(1)$$

We now use Double-E Rule together with relations (1) to obtain

$$\mathrm{E}(X^2Y^2) = \mathrm{E}\{\mathrm{E}(X^2Y^2 \mid y)\} = \mathrm{E}\{Y^2\mathrm{E}(X^2 \mid y)\}$$

$$= \mathrm{E}\{Y^2\,[\sigma_1^2(1-\rho^2) + (\rho^2\sigma_1^2/\sigma_2^2)Y^2]\}$$

$$= \sigma_1^2(1-\rho^2)\mathrm{E}(Y^2) + \rho^2(\sigma_1^2/\sigma_2^2)\mathrm{E}(Y^4) = \sigma_1^2\sigma_2^2(1-\rho^2) + 3\rho^2\sigma_1^2\sigma_2^2. \qquad \text{[by 1(d)]}$$

Thus $\mathrm{E}(X^2Y^2) = \sigma_1^2\sigma_2^2\,(1+2\rho^2)$ $\qquad \ldots(2)$

Since
$$\mathrm{E}(XY) = \mathrm{E}\{\mathrm{E}(XY \mid y)\} = \mathrm{E}\{Y\,\mathrm{E}(X \mid y)\} = \mathrm{E}(Y\,.\,(\rho\sigma_1/\sigma_2)Y\}$$

$$= (\rho\sigma_1/\sigma_2)\mathrm{E}(Y^2) = \rho\sigma_1\sigma_2$$

$$\therefore \quad \mathrm{Var}\,(XY) = \mathrm{E}(X^2Y^2) - \mathrm{E}^2(XY) = \sigma_1^2\sigma_2^2\,(1+\rho^2).$$

Now $\mathrm{Cov}\,(X^2, Y^2) = \mathrm{E}(X^2Y^2) - \mathrm{E}(X^2)\mathrm{E}(Y^2) = \sigma_1^2\,\sigma_2^2\,(1+2\rho^2) - \sigma_1^2\,\sigma_2^2 = 2\rho^2\,\sigma_1^2\sigma_2^2$

$$\mathrm{Var}\,(X^2) = \mathrm{E}(X^4) - \mathrm{E}^2(X^2) = (3\sigma_1^4) - (\sigma_1^2)^2 = 2\sigma_1^4; \ \mathrm{Var}\,(Y^4) = 2\sigma_2^4.$$

$$\therefore \quad \mathrm{Corr}\,(X^2, Y^2) = \frac{\mathrm{Cov}\,(X^2, Y^2)}{\sqrt{\mathrm{Var}\,X^2}\,\sqrt{\mathrm{Var}\,Y^2}} = \frac{2\rho^2\sigma_1^2\,\sigma_2^2}{2\sigma_1^2\,\sigma_2^2} = \rho^2.$$

Note. $\mathrm{E}(X^2, Y^2)$ can be had through m.g.f. See Example 13.

17*. Let us integrate **out** y from (1); then

$$\int_{-\infty}^{\infty} f(x, y)\,dy = \frac{1}{4\pi\sqrt{1-\rho^2}}\left\{\int_{-\infty}^{\infty} e^{-(1/2)Q(\rho)}dy + \int_{-\infty}^{\infty} e^{-(1/2)Q(-\rho)}dy\right\}$$

$$= \frac{1}{2\pi\sqrt{1-\rho^2}}\int_{-\infty}^{\infty} e^{-(1/2)Q(\rho)}\,dy$$

[Since changing y to $-y$, makes 2nd integral equal to 1st integral]. Put $y - \rho x = t\sqrt{1-\rho^2}$, then

$$Q(\rho) = [(y-\rho x)^2/(1-\rho^2)] + x^2 = t^2 + x^2, \quad dy = \sqrt{1-\rho^2}\,dt$$

$$\therefore \quad \int_{-\infty}^{\infty} f(x, y)\,dy = \frac{e^{-x^2/2}}{\sqrt{2\pi}}\left(\int_{-\infty}^{\infty}\frac{e^{-t^2/2}}{\sqrt{2\pi}}\,dt\right) = \frac{e^{-x^2/2}}{\sqrt{2\pi}}. \qquad \ldots(2)$$

From (2) we infer instantly that $\iint f(x, y)\,dx\,dy = 1$. Since $f(x, y) \geq 0 \ \forall \ x, y \in \mathbf{R}$, it follows that $f(x, y)$ is a joint density function. Incidently (2) yields that $X \sim N(0, 1)$. By symmetry of $f(x, y)$, we have $Y \sim N(0, 1)$. As $f_1(x)\,f_2(y) \neq f(x, y)$, X and Y are not independent.

$$\mathsf{E}(XY) = (4\pi\sqrt{1-\rho^2})^{-1}\left\{\iint xy\exp[-\tfrac{1}{2}Q(\rho)]\,dx\,dy + \iint xy\exp[-\tfrac{1}{2}Q(-\rho)]\,dx\,dy\right\}$$

Change y to $-y$ in the 2nd integral, the above Eq. reduces to

$$\mathsf{E}(XY) = (4\pi\sqrt{1-\rho^2})^{-1}\left\{\iint xy\exp[-\tfrac{1}{2}Q(\rho)]\,dx\,dy - \iint xy\exp[-\tfrac{1}{2}Q(\rho)]\,dx\,dy\right\} = 0.$$

$\therefore \quad \text{Cov}(X, Y) = \mathsf{E}(XY) - \mathsf{E}(X)\,\mathsf{E}(Y) = 0 \;\Rightarrow\; \text{Corr}(X, Y) = 0.$

If $(X, Y) \sim$ BVN (... ;), $\rho = 0 \Leftrightarrow$ independence, which is not true here, Hence (X, Y) is not BVN.

18*. Let $f(x, y) = Ae^{-(1+x^2)(1+y^2)}, \; -\infty < x, \; y < \infty.$ [A is Norming const.]

$$f_1(x) = Ae^{-(1+x^2)}\int_{-\infty}^{\infty} e^{-(1+x^2)y^2}\,dy = Ae^{-(1+x^2)}[\pi/(1+x^2)]^{1/2} \quad \left\{\because \int_{\infty}^{\infty} e^{-\lambda z^2}\,dz = \sqrt{\pi/\lambda}\right\}$$

$$f_2(y \mid x) = \frac{f(x, y)}{f_1(x)} = \frac{e^{-(1+x^2)y^2}}{\sqrt{\pi/(1+x^2)}} = \frac{\exp(-y^2/2\sigma_X^2)}{\sqrt{2\pi}\,\sigma_X}, \quad \{\sigma_X^2 = [2(1+x^2)]^{-1}\}$$

So $(Y \mid x) \sim N(0, \sigma_X^2)$. Similarly, $(X \mid y) \sim N(0, \sigma_Y^2)$ where $\sigma_Y^2 = [2(1+y^2)]^{-1}$.

Thus, conditional densities are normal, though $f(x, y)$ is non-Gaussian.

19*. Recall : $\max(X, Y) = \frac{1}{2}(X+Y) + \frac{1}{2}|X-Y|$

Let $Z = X - Y$; then $\mathsf{E}(Z) = 0$, Var $(Z) = 1 + 1 - 2\rho = 2(1-\rho) = \sigma^2$, say. Thus Z is $N(0, \sigma^2)$.

$$\therefore \quad f(z) = (1/\sqrt{2\pi}\,\sigma)\,e^{-(z^2/2\sigma^2)}, \quad -\infty < z < \infty.$$

$$\mathsf{E}\,|Z| = \int_{-\infty}^{\infty} |z|\,f(z)\,dz = 2\int_0^{\infty} zf(z)\,dz = \sigma\sqrt{2/\pi}\int_0^{\infty} e^{-u}\,du = \sigma\sqrt{2/\pi}. \quad [z^2 = 2u\sigma^2]$$

$$\therefore \quad \mathsf{E}[\max(X, Y)] = \tfrac{1}{2}\mathsf{E}(X+Y) + \tfrac{1}{2}\mathsf{E}|Z| = 0 + \frac{\sigma}{\sqrt{2\pi}} = \left(\frac{1-\rho}{\pi}\right)^{1/2}$$

Further, $\mathsf{E}\{\max(X, Y) + \min(X, Y)\} = \mathsf{E}(X + Y) = 0$. Thus

$$\mathsf{E}[\min(X, Y)] = -\,\mathsf{E}[\max(X, Y)] = -[(1-\rho)/\pi]^{1/2}.$$

Chapter 21 : Chi-Square Distribution

Sec. 21-42. Page 625

1*. The p.d.f. of $X \sim \chi^2_{(n)}$ is $f(x) = Ke^{-x/2}.\,x^{(n/2)-1}, \quad \{K = [(2)^{n/2}\,\Gamma(\tfrac{1}{2}n)]^{-1}\}.$

$$\therefore \quad P_0 = P\{X > x_0\} = \int_{x_0}^{\infty} \frac{1}{2}e^{-x/2}dx = e^{-x_0/2} \qquad [\text{When } n = 2,\ K = 1/2]$$

This result is equivalent to $x_0 = -2\ln P_0 \Rightarrow \chi_0^2 = 2\ln(1/P_0)$. When $P_0 = 0.05$, this gives

$$\chi_0^2 = 2\ln_e(20) = 2(2.9957322) = 5.9914645.$$

2*. Here $y = f(x) = \lambda^2 x e^{-\lambda x}$, $0 < x < \infty$. Differentiations yield

$f'(x) = \lambda^2(1-\lambda x)e^{-\lambda x}$, $f''(x) = \lambda^3(\lambda x - 2)e^{-\lambda x}$, $f'(x) = 0 \Rightarrow x = 1/\lambda$, $f''(\lambda^{-1}) = -\lambda^3/e < 0$.

Thus, $x = 1/\lambda$ is the mode and since $X = 2$ is the unique mode provided, we must have $\lambda = \frac{1}{2}$. So $f(x) = \frac{1}{4}e^{-x/2}x$, $x > 0$ and this incidently is $\chi^2_{(4)}$. It now follows that

$$P\{X < 9.49\} = P\{\chi^2_{(4)} < 9.49\} = 0.95, \text{ (from Tables)}.$$

3*. We use the m.g.f. of χ^2_n and p.g.f. of geom (p). Now by Double-E Rule.

$$\mathrm{E}(e^{tX}) = \mathrm{E}_N\{\mathrm{E}(e^{tX} \mid N = n)\} = \mathrm{E}_N[(1-2t)^{-2N/2}] = \mathrm{E}(\theta)^N. \quad [\theta = (1-2t)^{-1}]$$

$$= p/(1-q\theta) = p(1-2t)/(p-2t).$$

Sec. 21-72. Page 629

1*. Here $\sigma = 1$; $Y = [(X_{n+1} - \mu)/\sigma]^2$ is $[N(0, 1)]^2$, so that $Y \sim \chi^2_{(1)}$. Further, $Z = \dfrac{nS^2}{\sigma^2} = \sum_{i=1}^{n}(X_i - \overline{X})^2$ is $\chi^2_{(n-1)}$. Obviously, Y and Z are independent. Hence, by Reproductive Property

$$V = (Y + Z) \sim \chi^2_{(n)}.$$

2*. By definition : $(n-1)\hat{S}^2 = \Sigma(X_i - \overline{X})^2 = \Sigma[(X_i - \mu) - (\overline{X} - \mu)]^2$

$$= \Sigma(X_i - \mu)^2 + n(\overline{X} - \mu)^2 - 2n(\overline{X} - \mu)^2 = \Sigma(X_i - \mu)^2 - n(\overline{X} - \mu)^2$$

$$\therefore \quad (n-1)\mathrm{E}(\hat{S}^2) = \Sigma\mathrm{E}(X_i - \mu)^2 - n\mathrm{E}(\overline{X} - \mu)^2 = n\sigma^2 - n(\sigma^2/n) = (n-1)\sigma^2.$$

Thus, $\mathrm{E}(\hat{S}^2) = \sigma^2$. We have used Lin E in above steps.

When X is $N(\mu, \sigma^2)$, we know that $Y = (n-1)\hat{S}^2/\sigma^2 \sim \chi^2_{(n-1)}$. Thus $\hat{S} = \sigma\sqrt{Y}/\sqrt{n-1}$.

$$\therefore \quad \mathrm{E}(\hat{S}) = \frac{\sigma}{\sqrt{n-1}}\mathrm{E}(\sqrt{Y}) = \frac{\sigma}{\sqrt{n-1}}\int_0^\infty \frac{e^{-y/2} \cdot y^{(n/2)-1}}{\Gamma[(n-1)/2] \cdot 2^{(n-1)/2}}\,dy$$

$$= \frac{\sigma}{\sqrt{n-1}} \cdot \frac{1}{\Gamma[\frac{1}{2}(n-1)]} \frac{1}{(2)^{(n-1)/2}} \frac{\Gamma(\frac{1}{2}n)}{(\frac{1}{2})^{n/2}} = \sigma\frac{\sqrt{2}}{\sqrt{n-1}} \frac{\Gamma(\frac{1}{2}n)}{\Gamma[\frac{1}{2}(n-1)]}.$$

Obviously, $\mathrm{E}(\hat{S}) \neq \sigma$.

3*. The p.d.f. of $\hat{S}_1^2 = X$ (say) is given by [§ 21-72]

$$dF(x) = \left(\frac{n_1 - 1}{2\sigma_1^2}\right)^{(n_1-1)/2} \frac{x^{(n_1-3)/2}\exp[-(n_1-1)x/2\sigma_1^2]\,dx}{\Gamma[(n_1-1)/2]}, \quad 0 < x < \infty.$$

To shorten writing, let $(n_1 - 1)/2 = a$, $(n_2 - 1)/2 = b$, $(n_3 - 1)/2 = c$. The joint distribution of S_1^2, S_2^2, S_3^2 for $0 < x, y, z < \infty$ is

$$f(x, y, z) = \left(\frac{a}{\sigma_1^2}\right)^a \cdot \left(\frac{n}{\sigma_2^2}\right)^b \cdot \left(\frac{c}{\sigma_3^2}\right)^c \cdot \frac{x^{a-1}\, y^{b-1}\, z^{c-1} \exp\{-[(ax/\sigma_1^2) + (by/\sigma_2^2) + (cz/\sigma_3^2)]\}}{\Gamma(a)\,\Gamma(b)\,\Gamma(c)}$$

Let $u = x/z$, $v = y/z$, $w = z$, so that

$$J^{-1} = \frac{\partial(u, v, w)}{\partial(x, y, z)} = \begin{vmatrix} 1/z & 0 & -x/z^2 \\ 0 & 1/z & -y/z^2 \\ 0. & 0 & 1 \end{vmatrix} = \frac{1}{z^2}\,; dx\,dy\,dz = w^2\, du\,dv\,dw.$$

The joint distribution of U, V, W is, thus

$$f(u, v, w) = Ku^{a-1}\, v^{b-1}\, w^{a+b+c-1} \exp\{-[au/\sigma_1^2) + (bv/\sigma_2^2) + (c/\sigma_3^2)]\, w\}. \qquad \ldots(1)$$

where $K = \left(\frac{a}{\sigma_1^2}\right)^a \cdot \left(\frac{b}{\sigma_2^2}\right)^b \cdot \left(\frac{c}{\sigma_3^2}\right)^c \Gamma(a)\,\Gamma(b)\,\Gamma(c)$.

To obtain the joint density of U and V, we integrate **out** the unwanted variate W, $(0 < w < \infty)$.

$$f(u, v) = Ku^{a-1}\, v^{b-1} \int_0^\infty w^{a+b+c-1} \exp\left[-\left(\frac{au}{\sigma_1^2} + \frac{bv}{\sigma_2^2} + \frac{c}{\sigma_3^2}\right) w\right] dw$$

$$= \frac{K.\Gamma(a+b+c)\, u^{a-1}\, v^{b-1}}{[(au/\sigma_1^2) + (bv/\sigma_2^2) + (c/\sigma_3^2)]^{a+b+c}}, \quad 0 < u, \quad v < \infty$$

$$= \frac{a^a\, b^b\, c^c}{\sigma_1^{2a} . \sigma_2^{2b} . \sigma_3^{2c}} \cdot \frac{\Gamma(a+b+c)}{\Gamma(a)\,\Gamma(b)\,\Gamma(c)} \cdot \frac{u^{a-1}\, v^{b-1}}{[(au/\sigma_1^2) + (bv/\sigma_2^2) + (c/\sigma_3^2)]^{a+b+c}}, \quad 0 < u, v < \infty,$$

where a, b, c are as defined above.

Note. Use $\int_0^\infty \frac{x^{m-1}\, dx}{(A+Cx)^{m+n}} = \frac{B(m, n)}{A^n\, C^m}$, to verify normality.

Sec. 21-91. Page 635

1*. $M[t : \chi^2_{(n)}] = (1 - 2t)^{-n/2}$.

Letting $n \to 0$, this gives $M(t : \chi^2_{(0)}) = 1 = \Sigma p_i e^{t.0}$ which shows that χ_0^2 is degenerate variate :

$$P\{\chi^2_{(0)} = 0\} = 1.$$

2*. $M(t : Y) = \mathrm{E}(e^{tY}) = \mathrm{E}(e^{t \ln X}) = \mathrm{E}(X^t)$

$$= \int_0^\infty \frac{e^{-x/2}\, x^{(n/2)} . x^t}{\Gamma(\frac{1}{2}n)\,(2)^{n/2}}\, dx = \frac{\Gamma(\frac{1}{2}n + t)}{\Gamma(\frac{1}{2}n)\,(2)^{n/2}\,(\frac{1}{2})^{t+n/2}} = \frac{2^t\,\Gamma(\frac{1}{2}n + t)}{\Gamma(\frac{1}{2}n)}. \qquad \ldots(1)$$

$E(F^k) = \mathrm{E}(X_1^k . X_2^{-k}) = \mathrm{E}(X_i^k)\,\mathrm{E}(X_2^{-k}) \qquad [\because X_i \text{ are indep.}]$

$$= \frac{2^k\,\Gamma(\frac{1}{2}n + k)}{\Gamma(\frac{1}{2}n)} \cdot \frac{2^{-k}\,\Gamma(\frac{1}{2}n - k)}{\Gamma(\frac{1}{2}n)} = \frac{\Gamma(\frac{1}{2}n + k)\,\Gamma(\frac{1}{2}n - k)}{[\Gamma(\frac{1}{2}n)]^2}. \qquad \text{[by (1)]}$$

3*. Recall that, $Y = nS^2/\sigma^2$ is $\chi^2_{(n-1)}$. [§21-70] Hence

$$E(nS^2/\sigma^2) = n-1, \quad \text{Var}(nS^2/\sigma^2) = 2(n-1).$$

Thus, $$E(S^2) = (n-1)\sigma^2/n, \quad \text{Var}(S^2) = 2(n-1)\sigma^4/n^2.$$

Chapter 22 : Fisher-Student t-Distribution. Snedecor-Fisher F-Distribution

Sec. 22-34. Page 646

1*. (i) Here C is the normalizing (or norming) constant. Now

$$E(X^k) = C\int_{-a}^{a} x^k \left(1-\frac{x^2}{a^2}\right)^m dx.$$

For odd-valued k, integrand is an odd function in $(-a, a)$ and so integral is zero. Thus $E(X) = 0$. Now

$$E(X^{2k}) = 2C\int_0^a x^{2k}\left(1-\frac{x^2}{a^2}\right)^m dx = Ca^{2k+1}\int_0^1 z^{k+(1/2)-1}(1-z)^m\, dz, \qquad \left[\frac{x^2}{a^2}=z\right]$$

$$= Ca^{2k+1}\, B(k+\tfrac{1}{2}, m+1).$$

Put $k = 0$ to get $1 = CaB(1/2, m+1) \Rightarrow C = 1/a\, B(1/2, m+1)$. ...(1)

$$\therefore \quad \mu_{2k} = E(X^{2k}) = a^{2k}\frac{B(k+\frac{1}{2}, m+1)}{B(\frac{1}{2}, m+1)} = a^{2k}\frac{\Gamma(k+\frac{1}{2})}{\Gamma(\frac{1}{2})}\frac{\Gamma(m+3/2)}{\Gamma(k+m+3/2)} = \frac{a^{2k}.(\frac{1}{2})^{[k]}}{(m+3/2)^{[k]}}.$$

So $\mu_2 = E(X^2) = a^2/(2m+3) \Rightarrow a^2 = (2m+2)\mu_2$

$$\mu_4 = E(X^4) = \frac{3a^4}{(2m+3)(2m+5)};\ \beta_2 = \frac{\mu_4}{\mu_2^2} + \frac{3(2m+3)}{(2m+5)} \Rightarrow m = \frac{9-\beta_2}{2(\beta_2-3)}.$$

(ii) ***Relation between X and T.*** Using $2(m+1) = n$, it is

$$\frac{x}{a} = \left(1+\frac{n}{t^2}\right)^{-1/2};\ \frac{1}{a}\frac{dx}{dt} = \frac{1}{\sqrt{n}}\left(1+\frac{t^2}{n}\right)^{-3/2};\ 1-\frac{x^2}{a^2} = \frac{n}{n+t^2} = \left(1+\frac{t^2}{n}\right)^{-1} \qquad ...(2)$$

$$dF_T(t) = C\left(1+\frac{t^2}{n}\right)^{-m}\frac{a}{\sqrt{n}}\left(1+\frac{t^2}{n}\right)^{-3/2} dt = \frac{1}{\sqrt{n}}\frac{(1+t^2/n)^{-(n+1)/2}}{B(\frac{1}{2}, \frac{1}{2}n)}$$

Thus, $T \sim t_{(n)}$, where $n = 2(m+1)$.

Special Cases. From (2), $t^2 = nx^2/(a^2 - x^2)$, so at $n = 2$, $(m = 0)$

$$P\{T_2 > 2\} = P\left\{X^2 > \frac{4a^2}{n+4}\right\} = P\left\{X > \frac{\sqrt{2}}{3}a\right\} = C\int_{a\sqrt{2/3}}^{a} dx = Ca(1-\sqrt{2/3})$$

$$= (\sqrt{3}-\sqrt{2})/2\sqrt{3} = (\tfrac{1}{2}) - (5\sqrt{2}/16) \qquad \text{[Using (1) with } m = 0\text{]}$$

When $n = 4$, $(m = 1)$.

$$P\{T_4 > 2\} = P\left\{X > \left(\frac{a}{\sqrt{2}}\right)\right\} = \int_{a/\sqrt{2}}^{a} C\left(1-\frac{x^2}{a^2}\right)dx = aC\left(\frac{2}{3}-\frac{5}{6\sqrt{2}}\right) = \frac{1}{2}-\frac{5\sqrt{6}}{16}.$$

2*. $\overline{X} \sim N(\mu, \sigma^2/n), X' \sim N(\mu, \sigma^2)$, hence $(X'-\overline{X}) \sim N[0, \sigma^2 + (\sigma^2/n)]$ which provides

$$Z = \frac{(X'-X)}{\sigma\sqrt{(n+1)/n}} = \frac{(X'-X)\sqrt{n}}{\sigma\sqrt{(n+1)}} \sim N(0,1)$$

Also $(nS^2/\sigma^2) = \Sigma(X_j - \overline{X})^2/\sigma^2 = (S_0^2/\sigma^2) \sim \chi^2_{(n-1)}$

$$\therefore \qquad T = \frac{Z}{[\chi^2_{(n-1)}/(n-1)]^{1/2}} = \frac{(X'-\overline{X})\sqrt{n(n-1)}}{S_0\sqrt{n+1}} \sim t_{(n-1)}.$$

3*. Recall :

$$\mathrm{E}(X^k) = 2^k\,\Gamma(\tfrac{1}{2}n+k)/\Gamma(\tfrac{1}{2}n) \qquad [\S\ 21\text{-}20] \qquad ...(1)$$

$$\mathrm{E}(t^{2r}) = n^r\frac{\Gamma(\frac{1}{2}+r)}{\Gamma(\frac{1}{2})}\frac{\Gamma(\frac{1}{2}n-r)}{\Gamma(\frac{1}{2}n)}. \qquad [\S\ 22\text{-}21] \qquad ...(2)$$

Since $\dfrac{(n-1)S^2}{\sigma^2} = Z \sim \chi^2_{(n-1)}$ so put $S = \dfrac{\sigma}{\sqrt{n-1}}Z^{1/2}$ and thus

$$\mathrm{E}(S^r) = \frac{\sigma^r}{(n-1)^{r/2}}\mathrm{E}(Z^{r/2}) = \frac{\sigma^r 2^{r/2}}{(n-1)^{r/2}}\frac{\Gamma[\frac{1}{2}(n-1+r)]}{\Gamma[\frac{1}{2}(n-1)]} \qquad [\text{by (1)}, n \to n-1] \quad ...(3)$$

Since $\mathrm{E}(T) = 0$, Var $(t_n) = n/(n-2)$, hence Var $(T) = (n-1)/(n-3)$. Now

$$\begin{aligned}\mathrm{Cov}(\overline{X}, T) &= \mathrm{E}(\overline{X}.T) = \mathrm{E}(\overline{X}-\mu)T] \qquad (\because \mathrm{E}(T)=0)\\ &= \mathrm{E}[(\overline{X}-\mu)^2\sqrt{n}/S] = \sqrt{n}\,\mathrm{E}(\overline{X}-\mu)^2\,\mathrm{E}(S^{-1}). \qquad [\because T = (\overline{X}-\mu)\sqrt{n}/S]\\ &= \sqrt{n}(\sigma^2/n)\frac{2^{-1/2}\sqrt{n-1}\,\Gamma(n-2)/2]}{\sigma\Gamma[(n-1)/2]} \qquad [\text{by (3)}]\end{aligned}$$

$$\mathrm{Corr}(\overline{X}, T) = \frac{\mathrm{Cov}(\overline{X}, T)}{\sqrt{\mathrm{Var}(\overline{X})\,\mathrm{Var}(T)}} = \left(\frac{n-3}{2}\right)^{1/2}\frac{\Gamma[\frac{1}{2}(n-2)]}{\Gamma[\frac{1}{2}(n-1)]}.$$

4*. Here, $f(x) = K[1+(x^2/n)]^{-(n+1)/2}$, $-\infty < x < \infty$, $K^{-1} = \sqrt{n}\,B(\frac{1}{2}n, \frac{1}{2})$.

Let $Z = (n-\frac{1}{2})\ln[1+(X^2/n)]$; we find m.g.f. for Z. Thus

$$\begin{aligned}M(t:Z) &= \mathrm{E}(Z) = \mathrm{E}\{\exp\ln[1+(X^2/n)]^{t(n-1/2)} = \mathrm{E}\{[1+(X^2/n)]^{t(n-1/2)}\\ &= K\int_{-\infty}^{\infty}\left(1+\frac{x^2}{n}\right)^{t(n-1/2)}\left(1+\frac{x^2}{n}\right)^{-(n+1)/2} dx. \qquad ...(1)\end{aligned}$$

Here the integrand is Even, so we halfen the range to $(0, \infty)$; also we put $[1 + (x^2/n) = 1/y$, so that $dx = [\sqrt{n}/2y^{3/2}(1-y)^{1/2}]dy$. If follows that (1) reduces to, putting for K.

$$M(t:Z) = \int_0^1 \frac{y^{a-1}(1-y)^{(1/2)-1}\,dy}{B(\frac{1}{2}n, \frac{1}{2})} \qquad [a = \tfrac{1}{2}n(1-2t) + \tfrac{1}{2}t]$$

$$= B(a, \tfrac{1}{2}) / B(\tfrac{1}{2}n, \tfrac{1}{2}) = \Gamma(a)\,\Gamma(\tfrac{1}{2}n + \tfrac{1}{2}) / \Gamma(a + \tfrac{1}{2})\,\Gamma(\tfrac{1}{2}n)$$

$$= \frac{(\frac{1}{2}n)^{1/2}\,[1 - (1/4n)]}{(a)^{1/2}\,[1 - (1/8a)]} \qquad \text{[by given approximation]}$$

$$= \left(1 - \frac{1}{4n}\right)\left[1 - 2t + \frac{t}{n}\right]^{-1/2}\left(1 - \frac{1}{8a}\right)^{-1}.$$

We let $n \to \infty$, $[a^{-1} \to 0]$ then this m.g.f. simplifies to

$M(t : Z) \to (1 - 2t)^{-1/2} = M(t : \chi^2_{(1)}) \Rightarrow$ Z is approximately $\chi^2_{(1)}$-distributed.

Sec. 22-82. Page 657

1*. The elemental probability differential for $X \sim B_{II}(a, b)$ is

$$dP_1(x) = [B(a, b)]^{-1}\, x^{a-1}\,(1-x)^{b-1}\,dx, 0 < x < 1. \qquad \text{...(1)}$$

Let us put $z = bx / a(1-x)$, then $x = az/(b + az)$ and thus

$$1 - x = b/(b+az),\ dx = ab\,dz / (b+az)^2.$$

The function $z = f(x)$ is an increasing function; $0 < x < 1 \Rightarrow z > 0$. Making substitutions into (1) we obtain

$$dP_2(z) = \frac{1}{B(a,b)}\left(\frac{az}{b+az}\right)^{a-1}\left(\frac{b}{b+az}\right)^{b-1}\frac{ab\,dz}{(b+az)^2} = \frac{a^a b^b}{B(a,b)}\cdot\frac{z^{a-1}dz}{(b+az)^{a+b}}, z > 0.$$

This shows that Z is $F(2a, 2b)$-distributed.

Remark. If X is $B_{II}(a, b)$, then $(1 - X)$ is $B_{II}(b, a)$ [Vide §19-25, Example 1]. Thus

$$\frac{1}{Z} = \frac{(1-X)/b}{X/a} = \frac{a(1-X)}{bX} = \frac{(1-X)/(1-\mu)}{X/\mu} \text{ is } F(2b, 2a) \text{ distributed.}$$

2*. (i) Since $X \sim t_{(n-2)}$, its p.d.f. [using $(n - 2) = m^2$] is

$$dP_1(x) = [mB(\tfrac{1}{2}, \tfrac{1}{2}m^2)]^{-1}\,[1 + (x^2/m^2)]^{-(n-1)/2},\ -\infty < x < \infty. \qquad \text{...(1)}$$

As suggested in the problem put $x = \frac{1}{2}m(1-y)y^{-1/2}$, then

$$dx = -\tfrac{1}{4}m[(1+y)/y^{3/2}]\,dy,\ 1 + (x^2/m^2) = (1+y)^2/4y,$$

Making substitutions into (1), using $B(k, k) = 2^{1-2k}B(\frac{1}{2}, k)$, we get

$$dP_2(y) = \frac{1}{4B(\frac{1}{2}, \frac{1}{2}m^2)}\frac{1+y}{y^{3/2}}\cdot\frac{(4y)^{(n-1)/2}}{(1+y)^{n-1}}dy = \frac{2^{n-3}}{B(\frac{1}{2}, \frac{1}{2}m^2)}\frac{y^{(n/2)-2}dy}{(1+y)^{n-2}}, y > 0$$

$$= \frac{y^{(m^2/2)-1}}{B(\frac{1}{2}m^2, \frac{1}{2}m^2)(1+y)^{(m^2+m^2)/2}}.$$

This shows that $Y \sim F(m^2, m^2)$; $m^2 = n - 2$.

(ii) Here $\quad dF_X(x) = x^{(n/2)-1} dx / B(\frac{1}{2}n, \frac{1}{2}m)\,(1+x)^n, \; 0 < x < 1$ ...(1)

Now, $\quad B(k, k) = (2)^{1-2k} B(\frac{1}{2}, k) \Rightarrow B(\frac{1}{2}n, \frac{1}{2}n) = B(\frac{1}{2}, \frac{1}{2}n)/2^{n-1}$.

$$y = \frac{1}{2}\sqrt{n}\left(\frac{x-1}{\sqrt{x}}\right) \Rightarrow 1 + \frac{y^2}{n} = \frac{(x+1)^2}{4x} \Rightarrow \left(1 + \frac{y^2}{n}\right)^{1/2} = \frac{x+1}{2\sqrt{x}}.$$

Also, $$\frac{dy}{dx} = \frac{\sqrt{n}}{4x^{3/2}}(1+x) = \frac{\sqrt{n}}{2x}\left(1 + \frac{y^2}{n}\right)^{1/2}.$$

Making substitutions into (1), we get

$$dF_Y(y) = \frac{2^{n-1}}{B(\frac{1}{2}, \frac{1}{2}n)}\left(\frac{\sqrt{x}}{1+x}\right)^n \frac{dx}{x} = \frac{1}{\sqrt{n}} \frac{1}{B(\frac{1}{2}, \frac{1}{2}n)}\left(1 + \frac{y^2}{n}\right)^{-(n+1)/2} dy; \; -\infty < y < \infty.$$

This shows that $Y \sim t_{(n)}$.

3*. (i) If $X \sim F(m, n)$ and $Y \sim F(n, m)$, then for any $a > 0$, $P\{X > a\} + P(Y > 1/a) = 1$.
Take $a = 2.753$, $X \sim F(10, 12)$, $Y \sim F(12, 10)$; then $p_1 = P\{Y > 1/a\} = 1 - 0.05 = 0.95$.

(ii) Since $t_{(n)}^2 \sim F(1, n)$, we have

$$p_2 = P\{-\sqrt{4.747} < t_{12} < \sqrt{4.747}\} = P\{t_{12}^2 < 4.747\} = P\{F(1, 12) < 4.747\} = 1 - 0.05 = 0.95.$$

4*. For $X \sim F(2, n)$ the p.d.f. is

$$f(x) = 1/[1 + (2x/n)]^{(n/2)+1}, \; 0 < x < \infty, \; [\text{const.} = 1].$$

$$\therefore \quad P\{X \ge k\} = \int_k^\infty \left(1 + \frac{2x}{n}\right)^{-n/2-1} dx = \left\{\left(1 + \frac{2x}{n}\right)^{-n/2}\right\}_\infty^k = \left[1 + \frac{2k}{n}\right]^{-n/2}.$$

If $p = P(X \ge F)$, then this result gives $p = [1 + 2F/n)]^{-n/2}$. Thus $p^{-2/n} - 1 = 2F/n$ whence $F = \frac{1}{2}n[p^{-2/n} - 1]$: the significance level.

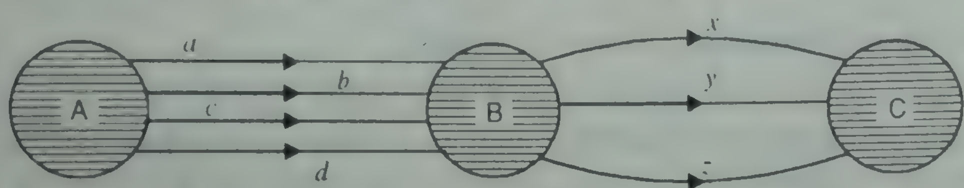
***Numbers don't lie, and they don't forgive.* (John Updike)**

Anyone, who lives within one's means,
suffers from a lack of Imagination. **(Oscar Wilde)**

Some Useful Formulas

1. *Basic Principle of Fundamental (or Sequential Counting) or Product Rule*

If a complex experiment can result in n_1 outcomes on the first stage, n_2 outcomes on the second stage, ..., n_k outcomes on the kth (i.e. last) stage, then the total number of outcomes for the complex experiment is $n_1 \,.\, n_2, \ldots.\, n_k$.



Example. If there are four routes joining the places A and B and three routes joining the places B and C, then there are $4 \times 3 = 12$ routes joining the places A and C via B, viz. *ax, ay, az, ..., dy, dz.*

Basic Rule of Sum

If an experiment E_1 can result in n_1 outcomes, a second experiment E_2 can result in n_2 outcomes, ..., kth distinct experiment E_k can result in n_k outcomes then the experiment E_1 or E_2 or, ..., or E_k will result in $n_1 + n_2 + \ldots + n_k$ outcomes, provided $E_1, E_2, \ldots, E_k$, cannot occur simultaneously.

Note. When the English word concerning numbers is or (in the exclusive sense : not both) use *Rule of Sum.* When the English word concerning numbers is *and*, use *Rule of Product.*

2. *Sampling of Two objects out of Four Distinct objects, Generalizations*

	x	y	z	t
x	-	xy	xz	xt
y	yx	-	yz	yt
z	zx	zy	-	zt
t	tx	ty	tz	-

Without replacement, with order : ${}^4P_2$

	x	y	z	t
x	-	xy	xz	xt
y	-	-	yz	yt
z	-	-	-	zt
t	-	-	-	-

Without replacement, without order : $\binom{4}{2}$

	x	y	z	t
x	xx	xy	xz	xt
y	yx	yy	yz	yt
z	zx	zy	zz	zt
t	tx	ty	tz	tt

With replacement, with order : 2^4

	x	y	z	t
x	xx	xy	xz	xt
y	-	yy	yz	yt
z	-	-	zz	zt
t	-	-	-	tt

With replacement, without order : $\binom{4+2-1}{2}$

A. Sampling without replacement but with order

A particular selection of r objects from n distinct objects, without replacement, but with regard to the order of selection, is called a *permutation* of n objects taken r at a time, is denoted by ^nP_r and is given by

$$^nP_r = n(n-1)(n-2)\dots(n-r+1) = n^{(r)}.$$

B. Sampling without replacement and without order

A particular selection of r objects from n objects, without replacement or any regard to order of selection, is called a *combination* of n objects taken r at a time, denoted by $\binom{n}{r}$, ^nC_r or $C(n, r)$ and is given by

$$^nC_r \equiv \binom{n}{r} = \frac{n!}{r!(n-r)!} = \frac{n(n-1)(n-2)\dots(n-r+1)}{r!} = \frac{n^{(r)}}{r!}. \qquad \text{[Binomial Coefficient]}$$

C. Sampling with replacement, but with order

A particular selection of r objects from n objects, with replacement and with regard to order of selection, is called *permutation with repetitions* of n objects taken r at a time, or accumulation of n objects taken r at a time. This may be denoted by $A(n, r)$ and is given by

$$A(n, r) = n^r.$$

D. Sampling with replacement and without order

A particular selection of r objects from n objects, with replacement and without any regard to order of selection is given by $\binom{n+r-1}{r}$.

Number of possible arrangements of size r from n objects

	Without replacement	*With replacement*
Ordered	$^nP_r = \dfrac{n!}{(n-r)!}$	n^r
Unordered	$\binom{n}{r} = \dfrac{n!}{(n-r)!\, r!}$	$\binom{n+r-1}{r}$

E. Theorem on grouped permutations

If n objects can be divided into k groups such that objects belonging to the same group are alike, while objects belonging to different groups are different, then the number of permutations of these objects taken all at a time, is

$$\binom{n}{n_1, n_2, \dots, n_k} = \frac{n!}{n_1!\, n_2!\, \dots n_k!}. \quad (n_1 + n_2 + \dots + n_k = n). \quad \text{[Multinomial coefficient]}$$

3. Occupancy Problems

The number of ways, b balls can be distributed among N distinguishable cells, depends upon the fact ; how many balls are permitted to be in one cell and whether the balls

are distinguishable from one another or not. Problems involving the distribution of balls among cells are called *occupancy problems*.

***Theorem* 1.** The number of ways in which b distinct balls can be alloted to N $(N \geq b)$ distinguishable cells, if at most one ball is permitted in a cell, is ${}^{N}P_{b} = N!/(N-b)!$

Proof. Since each cell receives either 0 or 1 ball, it follows that the number of ways that we can choose the cells to be assigned balls is ${}^{N}P_{b} = N!/(N-b)!$

***Theorem* 2.** The number of ways in which b distinguishable balls can be alloted to N distinguishable cells, if no restriction is placed on the No. of balls permitted in a cell, is N^{b}.

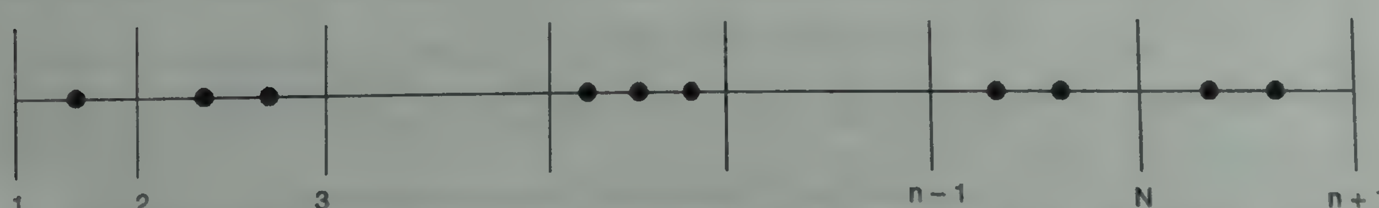
Proof. The first ball can be alloted in N ways. For each of these allotments, the second ball can also be allotted in N ways, etc. Thus, the total number of possible allotments is N^{b} (*Rule of product*).

Fermi-Dirac Formula. The number of ways b identical balls can be allotted to N distinguishable cells, if only one ball is permitted per cell, is $\binom{N}{b}$.

Proof. Since the balls are identical, the required No. of ways is, to choose b cells from N cells.

Bose-Einstein Formula. The number of ways b identical balls can be allotted to N labelled cells when no restriction is placed on the number of balls permitted in any cell is

$$\binom{N+b-1}{b} = \binom{N+b-1}{N-1}$$



Proof. Represent the N cells by the spaces between $N+1$ bars and represent the balls by • . Then there are $N+1+b$ positions filled with $N+1$ bars and b zeros. The first and last position will always be occupied by a bar. The required number is now the number of ways of placing b zeros in the remaining $N+b-1$ positions. This is given by

$\binom{N+b-1}{b}$ ways.

Integer Counting Rule. Find the number of solutions of the equation

$$x_1 + x_2 + \ldots + x_k = n. \qquad \ldots(1)$$

(i) when each x_i is strictly positive integer $(x_i > 0, \forall\ i)$

(ii) when each x_i is a non-negative integer $(x_i \geq 0, \forall\ i)$

Proof. (i) The number of solutions equals the number of ways of splitting a queue of n persons into k batches, each batch consisting of one or more persons.

Now when n persons are standing in a row, there are $(n-1)$ gaps between them.

A split of this queue into k batches can be had by inserting $(k-1)$ barriers into these $(n-1)$ gaps. This amounts to choosing $(k-1)$ slots out of $(n-1)$ slots. This objective is attained in $\binom{n-1}{k-1}$ ways. So, the numbers of positive integer solutions of equation (1) is

$$N_1 = \binom{n-1}{k-1}.$$

(ii) Write $x_i + 1 = y_i$, for $i = 1, 2, 3, \ldots, k$. Now $x_i \geq 0 \Rightarrow y_i \geq 1 \Rightarrow y_i > 0$. The Eq. (1) in terms of variables y_i is

$$y_1 + y_2 + \ldots + y_k = n + k, \qquad y_i > 0 \ \forall \ i$$

By part (i), the number of solutions must be

$$N_2 = \binom{n+k-1}{k-1}.$$

4. Matrix Inverse :

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{\Delta}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}, \quad \Delta = \det\begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc.$$

5. Factorial Powers

(i) The kth *decreasing* factorial power of x is denoted by $x^{(k)}$ and is defined by

$$x^{(k)} = x(x-1)(x-2)\ldots(x-k+1).$$

(ii) The kth *increasing* (or reverse) factorial power of x is denoted by $x^{[k]}$ and is defined by

$$x^{[k]} = x(x+1)(x+2)\ldots(x+k-1).$$

(iii) *Relations :* $x^{(k)} = (-1)^k (-x)^{[k]}$; $x^{[k]} = (-1)^k (-x)^{[k]}$; $N^{(k)} = N^{(k-x)}(N-k+x)^{(x)}$.

$$n! = n^{(r)}.(n-r)! \qquad n^{(r)} = (n-r+1)^{[r]}$$

$$\binom{n}{r} = \frac{n^{(r)}}{r!} = (-1)^r(-n)^{(r)}, \quad n^{[r]} = \binom{n+r-1}{r} r!; \quad x^{[r]}\binom{x-1}{k-1} = k^{[r]}\binom{x-1+r}{k-1+r}.$$

$$\binom{n}{r} = \frac{n^{(s)}}{r^{(s)}}\binom{n-s}{r-s} = \frac{(n-s+1)^{[s]}}{(r-s+1)^{[s]}}\binom{n-s}{r-s}. \qquad \left[\binom{n}{t} = 0, \text{ if } t > n\right]$$

$x^2 = x^{(2)} + x$; $x^3 = x^{(3)} + 3x^{(2)} + x$; $x^4 = x^{(4)} + 6x^{(3)} + 7x^{(2)} + x$; $x^2 = x^{[2]} - x$, $x^3 = x^{[3]} - 3x^{[2]} + x$. etc.

Note. Some authors write $(N)_x$ and $[N]^x$ for $N^{[x]}$.

$$\binom{n+1}{k} = \frac{n+1}{n+1-k}\binom{n}{k}, \quad \binom{n-1}{k} = \frac{n-k}{n}\binom{n}{k}, \quad \binom{n}{k+1} = \frac{n-k}{k+1}\binom{n}{k}, \quad \binom{n}{k-1} = \frac{k}{n-k+1}\binom{n}{k}$$

6. Fundamental Expansions

Taylor Series. The Taylor series for $f(x)$ about $x = a$ is given by

$$f(x) = f(a) + (x-a)f'(a) + \frac{(x-a)^2}{2!}f''(a) + \ldots + \frac{(x-a)^{n-1}}{(n-1)!}f^{n-1}(a) + R_n$$

where $f^k(a) = \left[d^k f(x)/dx^k\right]_{x=a}$; $R_n = f^n(\theta)(x-a)^n/n!,\ a \le \theta \le x.$

$$e^x = \exp(x) = 1 + x + \frac{x^2}{2!} + \ldots + \frac{x^n}{n!} + \ldots = \sum_{k=0}^{\infty} \frac{x^k}{k!}, \quad -\infty < x < \infty$$

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \ldots = \sum_{k=1}^{\infty} (-1)^{k-1} \frac{x^k}{k}$$

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \ldots + (-1)^{n-1} \frac{x^{2n-1}}{(2n-1)!} + \ldots = \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k+1}}{(2k+1)!}$$

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \ldots + (-1)^n \frac{x^{2n}}{2n!} + \ldots = \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k}}{(2k)!}$$

$$\tan x = x + \frac{x^3}{3} + \frac{2}{15} x^5 + \frac{17}{315} x^7 + \frac{62}{2835} x^9 + \ldots$$

$$\sin^{-1} x = x + \left(\frac{x^3}{3!}\right) + \left(\frac{9x^5}{5!}\right) + \ldots$$

$$\tan^{-1} x = x - \left(\frac{x^3}{3}\right) + \left(\frac{x^5}{5}\right) - \left(\frac{x^7}{7}\right) + \ldots = \sum_{i=0}^{n} \frac{(-1)^n x^{2n+1}}{(2n+1)}$$

$$\sin i\theta = i \sinh\theta, \quad \cos i\theta = \cosh\theta, \quad \cosh i\theta = \cos\theta, \quad \sinh i\theta = i \sin\theta$$

$$\sin\theta = \frac{e^{i\theta} - e^{-i\theta}}{2i}, \quad \cos\theta = \frac{e^{i\theta} + e^{-i\theta}}{2}; \quad \sinh\theta = \frac{e^{\theta} - e^{-\theta}}{2}; \cosh\theta = \frac{e^{\theta} + e^{-\theta}}{2}$$

Taylor series for $f(x, y)$ about $x = a$, $y = b$. Let $h = x - a$, $k = y - b$. Then

$$f(x, y) = f(a,b) + [hf_x(a,b) + kf_y(a,b)] + [h^2 f_{xx}(a,b) + 2hk\, f_{xy}(a,b) + k^2 f_{yy}(a,b)]/2! + \ldots$$

7. *Binomial Expansions and Binomial Coefficients*

(i) If n is a positive integer, then

$$(a+b)^n = \binom{n}{0} a^n + \binom{n}{1} a^{n-1} b + \ldots + \binom{n}{r} a^{n-r} b^r + \ldots + \binom{n}{n} b^n = \sum_{k=0}^{n} \binom{n}{r} a^{n-r} b^r$$

(ii) If n is any real number, then

$$(1+x)^n = 1 + nx + \frac{n(n-1)}{2!} x^2 + \ldots + \frac{n(n-1)(n-2)\ldots(n-r+1)}{r!} x^r + \ldots$$

Neg-bin Expansion

$$(1-x)^{-n} = \sum_{r=0}^{\infty} \binom{n+r-1}{r} x^r = \sum_{r=0}^{\infty} \binom{-n}{r} (-x)^r; \quad n > 0$$

$$\binom{n}{k} = \frac{n}{n-k} \binom{n-1}{k}; \qquad \binom{n}{k} = \frac{n-k+1}{k} \binom{n}{k-1}; \qquad \binom{n}{k} = \frac{n(k+1)}{(n-k)(n-k-1)} \binom{n-1}{k+1}.$$

These indicate head-decreased, tail-decreased or head-decreased-tail increased. For proofs, write, e.g. $\binom{n}{k} = \lambda\binom{n}{k-1}$ and find λ, etc. reversed results are immediate.

8. *The Hypergeometric Identity*

Let a, b be real numbers and n a positive integer. Then

$$\sum_{k=0}^{\infty}\binom{a}{k}\binom{b}{n-k} = \binom{a+b}{n} \qquad \text{...(A)}$$

Proof.

$$(1+y)^{a+b} = (1+y)^a(1+y)^b = \left[\sum_{k=0}^{a}\binom{a}{k}y^k\right]\left[\sum_{j=0}^{b}\binom{b}{j}y^j\right]$$

$$\therefore \quad \sum_{n=0}^{a+b}\binom{a+b}{n}y^n = \sum_{k=0}^{a}\sum_{j=0}^{b}\binom{a}{k}\binom{b}{j}y^{k+j}. \qquad \text{...(i)}$$

Equating Coefficient of $y^n\,[n = k+j]$ on both sides of (i) yields (A)

9. *A note on Geometric Progression*

$$1 + r + r^2 + \ldots + r^n = (1 - r^{n+1})/(1-r)$$

If $|a| < 1$, the infinite geometric series

$$\Sigma a^r = a + a^2 + a^3 + \ldots = a/(1-a), \qquad r = 1, 2, 3, \ldots \qquad \text{...(1)}$$

is uniformly convergent. Repeatedly differentiating (1) w.r.t. "a" we recover

$$\Sigma r a^{r-1} = 1/(1-a)^2 \quad \text{or} \quad \Sigma r a^r = a/(1-a)^{-2}. \qquad \text{...(2)}$$

$$\Sigma r^2 a^{r-1} = (1+a)(1-a)^{-3} \Rightarrow \Sigma r^2 a^r = a(1+a)(1-a)^{-3} \qquad \text{...(3)}$$

$$\Sigma r^3 a^{r-1} = [(1+a)^2 + 2a]/(1-a)^4 \Rightarrow \Sigma r^3 a^r = [2a^2 + a(1+a)^2]/(1-a)^4 \qquad \text{...(4)}$$

The process of differentiation can be carried out any number of times, etc.

Sums of Powers of Natural Numbers

$$\sum_{k=1}^{n} k = \frac{n(n+1)}{2}; \qquad \sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}; \qquad \sum_{k=1}^{n} k^3 = \frac{[n(n+1)]^2}{4};$$

$$\sum_{k=1}^{n} k^4 = \frac{n(n+1)(2n+1)(3n^2+3n-1)}{30}.$$

10. *Leibniz's Theorem on nth Derivative of a Product*

$$D^n(uv) = (D^n u).v + \binom{n}{1}(D^{n-1}u)(Dv) + \binom{n}{2}(D^{n-2}u)(D^2 v) + \ldots + \binom{n}{r}(D^{n-r}u)(D^r v) + \ldots + \binom{n}{n}(u)(D^n v).$$

11. *Some Properties of Definite Integrals*

$\int_{-a}^{a} f(x)\,dx = 0$, (if f is odd) ; $\int_{-a}^{a} f(x)\,dx = 2\int_0^a f(x)\,dx$, (if f is even).

$\int_0^{2a} f(x)\,dx = 0$, if $f(2a-x) = -f(x)$; $\int_0^{2a} f(x)\,dx = 2\int_0^a f(x)\,dx$, if $f(2a-x) = f(x)$.

Some Special Integrals

$$\int e^{ax}\sin bx\,dx = \frac{e^{ax}(a\sin bx - b\cos bx)}{(a^2+b^2)},\quad \int e^{ax}\cos bx\,dx = \frac{e^{ax}(a\cos bx + b\sin bx)}{a^2+b^2}.$$

$$\int \frac{dx}{a+b\cos x} = \frac{1}{\sqrt{a^2-b^2}}\cos^{-1}\left(\frac{b+a\cos x}{a+b\cos x}\right); a>b.\quad \int_0^\infty \frac{\cos mx\,dx}{b^2+x^2} = \frac{\pi}{2b}e^{-mb};\ m>0, b>0$$

$$\int_{-\infty}^{\infty} e^{-(ax^2+bx+c)}\,dx = \sqrt{\frac{\pi}{a}}\exp\left(\frac{b^2-4ac}{4a}\right),\quad a>0$$

$$\int_{-\infty}^{\infty} \exp\{-p^2x^2 \pm qx\}\,dx = \frac{\sqrt{\pi}}{p}\exp\left(\frac{q^2}{4p^2}\right),\quad p>0$$

$$\int_{-\infty}^{\infty} x^n e^{-px^2+2qx}\,dx = \frac{1}{2^{n-1}p}\sqrt{\frac{\pi}{p}}\frac{d^{n-1}}{dq^{n-1}}\left(qe^{q^2/p}\right),\quad p>0$$

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp\left[-\left(ax^2+2hxy+by^2\right)\right]dxdy = \frac{\pi}{\sqrt{ab-h^2}}.$$

$$\int_0^{\infty}\int_0^{\infty} \exp\left[-\left(ax^2+2hxy+by^2\right)\right]dxdy = \frac{\pi}{2\sqrt{ab-h^2}}\tan^{-1}\left(\frac{\sqrt{ab-h^2}}{h}\right).$$

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp\left[-\frac{1}{2}(ax^2+2hxy+by^2-2lx-2my\right]dx\,dy = \frac{2\pi}{\sqrt{ab-h^2}}\exp\left(\frac{1}{2}A\right),$$

where $A = (am^2-2hlm+bl^2)/(ab-h^2)$.

$$\int_\alpha^{\infty} \exp(-Ax^2-Bx)\,dx = \sqrt{\frac{\pi}{A}}\exp\left(\frac{B^2}{4A}\right)\left[1-\Phi\left(\frac{2A\alpha+B}{\sqrt{2A}}\right)\right],\ A>0$$

Proof. As $Ax^2+Bx = \left(x\sqrt{A}+B/2\sqrt{A}\right)^2 - (B^2/4A)$, we get

$$I = \exp\left(\frac{B^2}{4A}\right)\int_\alpha^{\infty}\exp\left[-\left(x\sqrt{A}+\frac{B}{2\sqrt{A}}\right)^2\right]dx. \qquad \left[\text{Put } x\sqrt{A}+\frac{B}{2\sqrt{A}} = \frac{t}{\sqrt{2}}\right]$$

$$= \exp\left(\frac{B^2}{4A}\right)\frac{\sqrt{\pi}}{\sqrt{A}}\left(\int_{t_0}^{\infty}\frac{e^{-\frac{1}{2}t^2}}{\sqrt{2\pi}}dt\right) = \sqrt{\frac{\pi}{A}}\exp\left(\frac{B^2}{4A}\right).[1-\Phi(t_0)], \qquad [\Phi(t_0): \text{c.d.f. of } N(0,1)]$$

$$= \sqrt{\frac{\pi}{A}}\exp\left(\frac{B^2}{4A}\right)\left[1-\Phi\left(\frac{2A\alpha+B}{\sqrt{2A}}\right)\right]$$

(Find values when $\alpha = 0$, $\alpha = -\infty$)

Kronecker General Rule of Integration by Parts

$$\int p(x)\, f(x)\, dx = pF_1 - p'F_2 + p''F_3 - p'''F_4 + \ldots + (-1)^m p^{(m)} F_{m+1}$$

where $p(x)$ is a polynomial of degree m, p', p'', ... are successive differentiations of p F_1 denotes an indefinite integral of f, F_2 an indefinite integral of F_1, and so on, and alternating signs are affixed to the terms. Use brackets for evaluations and simplify only when the process is complete.

12. Gamma Integral and Some of its Properties

$$\Gamma(a) = \int_0^\infty e^{-x} x^{a-1}\, dx, \quad a > 0 \qquad \text{(Definition)}$$

Working Definition : $\boxed{\int_0^\infty e^{-\lambda x} x^{a-1}\, dx = \dfrac{\Gamma(a)}{(\lambda)^a}, \quad \lambda > 0,\ a > 0.}$

(a) $\Gamma(k) = (k-1)\,!$, (if k is positive integer). (b) $\Gamma(-n) \to \infty$ ($n > 0$ is an integer)

(c) $\Gamma(x) = (x-1)\,\Gamma(x-1)$, [Recurrence Formula]. (d) $\Gamma(\frac{1}{2}) = \sqrt{\pi}$.

(e) $\Gamma(n+\frac{1}{2}) = 2n!\,\sqrt{\pi}/2^{2n}(n!)$. (f) $\sqrt{\pi}\,\Gamma(2n) = (2)^{2n-1}\Gamma(n)\,\Gamma(n+\frac{1}{2})$ (Duplication Formula).

(g) $\dfrac{\Gamma(a+r)}{\Gamma(a)} = a^{[r]}$; $\dfrac{\Gamma(b-r)}{\Gamma(b)} = \dfrac{1}{(b-r)^{[r]}}$; $\Gamma\left(\dfrac{n}{2}\right) = \dfrac{(n-1)!\,\sqrt{\pi}}{[(n-1)/2]!\,2^{n-1}}$ (n: odd)

(h) $\Gamma(p)\,\Gamma(1-p) = \dfrac{\pi}{\sin p\pi}$, $0 < p < 1$ (i) $\int_0^{\pi/2} \sin^p\theta \cos^q\theta\, d\theta = \dfrac{\Gamma[(p+1)/2]\,\Gamma[(q+1)/2]}{2\Gamma[(p+q+2)/2]}$.

13. Beta Integral

Let $a > 0$, $b > 0$.

$$B(a, b) = \int_0^1 x^{a-1}(1-x)^{b-1}\, dx; \quad B(a,b) = \int_0^\infty \frac{x^{a-1}\, dx}{(1+x)^{a+b}}; \quad \int_0^\infty \frac{x^{a-1}\, dx}{(1+kx)^c} = \frac{\Gamma(a)\,\Gamma(c-a)}{k^a\,\Gamma(c)},\ k > 0$$

$B(a, b) = B(b, a)$ (symmetry) ; $B(a, b) = \dfrac{\Gamma(a)\,\Gamma(b)}{\Gamma(a+b)}$. (Relation with Gamma integral)

Gamma Function Limit : $\lim\limits_{n\to\infty} \dfrac{\Gamma(n+k)}{(n)^k\,\Gamma(n)} = 1$, or $\dfrac{\Gamma(n+k)}{\Gamma(k)} \approx n^k$

Proof. Using gamma-beta relation we have

$$\lim_{n\to\infty} \frac{\Gamma(k)\,\Gamma(n)}{\Gamma(k+n)} n^k = \lim_{n\to\infty} n^k \int_0^1 x^{k-1}(1-x)^{n-1}\, dx \qquad \ldots(1)$$

Put $x = y/n$ in the R.H.S. to get

$$\lim_{n\to\infty} \int_0^n y^{k-1}\left(1-\frac{y}{n}\right)^{n-1} dy = \int_0^\infty y^{k-1} e^{-y}\, dy = \Gamma(k)$$

Thus (1) gives : $\lim\limits_{n\to\infty} \left\{\dfrac{n^k\,\Gamma(k)\,\Gamma(n)}{\Gamma(n+k)}\right\} = \Gamma(k) \Rightarrow \lim\limits_{n\to\infty} \left\{\dfrac{\Gamma(n+k)}{n^k\,\Gamma(n)}\right\} = 1.$

14. Euler's Limit

If $\lim \phi(n) = a$, as $n \to \infty$, then

$$\lim_{n\to\infty}\left(1+\frac{\phi(n)}{n}\right)^{bn+c} = e^{ab}; \quad \lim_{n\to\infty}\left(1+\frac{a}{n}\right)^{bn+c} = e^{ab}; \quad \lim_{n\to\infty}\left[1+\frac{a}{n}+O\left(\frac{1}{n}\right)\right]^{bn+c} = e^{ab}$$

15. Stirling's Approximation

When n is very large $\qquad n! \doteq \sqrt{2\pi}\, e^{-n}\,(n)^{n+1/2}$.

If n, k and $n-k$ are sufficiently large, this result provides

$$\binom{n}{k} = \frac{1}{\sqrt{2\pi n}}\left(\frac{n}{k}\right)^{k+1/2}\left(\frac{n}{n-k}\right)^{n-k+1/2}$$

16. Differentiation Under Integral Sign (DUIS)

Under differentiable conditions let

$$F(t) = \int_a^b f(t,x)\,dx; \ a, b \text{ functions of } t.$$

Then

$$\frac{dF(t)}{dt} = \int_a^b \frac{\partial f(t,x)}{\partial t}\,dx + f(t,b)\frac{db}{dt} - f(t,a)\frac{da}{dt}.$$

$$\boxed{\int_0^\infty \exp[-a^2x^2-(b^2/x^2)]\,dx = \frac{\sqrt{\pi}}{2a}e^{-2ab}; \ a>0,\ b\geq 0. \qquad \int_0^\infty e^{-\lambda x^2}\,dx = \frac{1}{2}\sqrt{\frac{\pi}{\lambda}}.}$$

17. Euler-Maclaurin Sum Formula

If $[x]$ = greatest integer contained in x, then

$$\sum_{k=0}^{\infty} f(k) = \int_0^n f(x)\,dx + \frac{1}{2}[f(0)+f(n)] + \int_0^n \left(x-[x]-\frac{1}{2}\right)f'(x)\,dx.$$

18. A note on Functions with Simple Discontinuities

Let a be any point within the domain of definition of a function f. Any point to the right of a is $x = a + \varepsilon$, ($\varepsilon > 0$). As $\varepsilon \to 0$ the point $a + \varepsilon \to a$ from the right. If the corresponding values $f(a + \varepsilon)$ tend to a limiting value then this value is called the *right hand limit* of f as $x \to a$ and is denoted by $f(a\,+)$. Formally

$$f(a+0) = \lim_{\varepsilon\to 0} f(a+\varepsilon).$$

If in addition, $\lim f(a+\varepsilon) = f(a)$, as $\varepsilon \to 0$, then f is said to be *continuous from the right at a. Similarly, f is said to be continuous from the left at a, if,*

$$f(a-0) = \lim_{\varepsilon\to 0} f(a+\varepsilon) = f(a).$$

Clearly, f is continuous at a iff it is continuous both from the right and from the left ; thus

$$\lim_{\varepsilon\to 0} f(a+\varepsilon) = f(a) = \lim_{\varepsilon\to 0} f(a+\varepsilon), \ (\varepsilon > 0).$$

If $f(a\,+)$ and $f(a\,-)$ both exist but are unequal in value, then f is said to have simple or *jump discontinuity* at a. The number k

$$k = f(a+0) - f(a-0) = f(a+) - f(a-) \qquad [\text{some write } k = f(a^+) - f(a^-)]$$

is called the *jump* or *saltus* of f at a.

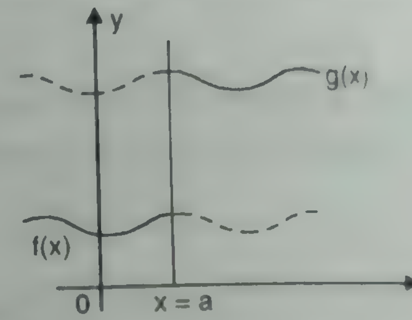
From elementary analysis, we borrow the simple results :

1. If f is monotonic on R, then $f(x+)$ and $f(x-)$ both exist for each $x \in R$. Further, both the infinite limits

$$\lim_{n \downarrow -\infty} f(x) = f(-\infty); \; \lim_{n \uparrow \infty} f(x) = f(+\infty)$$

also exist for monotonic function f.

2. A bounded monotonic function has atmost countable number of the points of jump discontinuities.



19. *Unit Step Function*

This is defined as under

$$u_c(x) = 1 \quad \text{for } x > 0$$
$$= c, \quad \text{for } x = 0$$
$$= 0, \quad \text{for } x < 0$$

Some authors define $u(0) = 0$, or $u(0) = 1$ or $u(0) = 1/2$ and each value has its defenders. Note that, $u(0) = 0$ gives a function u_0 which is continuous from the left at 0 ; $u(0) = 1$ gives a function u_1, which is continuous from the right at 0.

Importance of unit step function. Suppose ψ is a function continuous everywhere except at $x = a$, where it has a simple discontinuity. Then ψ can always be expressed as a linear combination of two functions f and g which are continuous everywhere but which have been truncated at $x = a$.

$$\psi(x) = f(x)\,u(a-x) + g(x)\,u(x-a) \text{ i.e. } \psi(x) = \begin{cases} f(x) & \text{for } x < a \\ g(x) & \text{for } x > a \end{cases}$$

20. *Derivative of the Unit Step Function. Dirac Delta Function*

For all $x \neq 0$, $u'(x) = 0$. At $x = 0$, however, there is a jump discontinuity and the definition of derivative accordingly fails. Further

$$\lim_{h \to 0} \left\{ \frac{u_c(h) - u_c(0)}{h} \right\} = +\infty.$$

Thus, from a descriptive point of view at least, the derivative of u would appear to be a function, called *Dirac delta function*, written $\delta(x)$, which has the following piece-wise specification :

$$\delta(x) \equiv u'(x) = \begin{cases} 0, & \text{for all } x \neq 0 \\ \infty, & \text{for } \quad x = 0 \end{cases}$$

21. *Sifting Property of Delta Function*

Let φ be any function continuous on a neighbourhood of zero, say for $-a < x < a$. Then we should have

$$\int_{-a}^{a} \varphi(x)\, u'(x)\, dx = \int_{-a}^{a} \varphi(x) . \lim_{h \to 0} \left(\frac{u(x+h) - u(x)}{h} \right) dx.$$

Assume that it is permissible to interchange the operations of integration and taking limits as $h \to 0$; this gives, using the first MVT of Integral Calculus,

$$\int_{-a}^{a} \varphi(x)\, u'(x)\, dx = \lim_{h \to 0} \frac{1}{h} \int_{-h}^{0} \varphi(x)\, dx = \lim_{h \to 0} \varphi(t)$$

where t is some point lying between $-h$ and 0. Since φ is assumed to be continuous in the neighbourhood of zero, it follows that

$$\int_{-a}^{a} \varphi(x)\, \delta(x)\, dx = \varphi(0) \qquad \text{[Sifting Property of Delta Function]}$$

$$\int_{c-\varepsilon}^{c+\varepsilon} \delta(x-c)\, f(x)\, dx = f(c),\ \forall f \text{ and } \forall \varepsilon > 0$$

Remarks : $\int_{-a}^{a} \varphi(x)\, \delta(x-c)\, dx = \varphi(c)$ [Sifting Property]

$$u'(x-a) = \delta(x-a),\ \delta(-x) = \delta(x). \quad \delta(x-a) = \begin{cases} 0, & \text{if } x \neq a \\ \infty, & \text{if } x = a \end{cases}$$

22. *Hyper-geometric Series*

A power series with parameters a, b, c and defined by

$$2F_1\left(\begin{matrix} a, b \\ c \end{matrix} \middle| x \right) = 1 + \frac{a.b}{1.c} x + \frac{a(a+1).b(b+1)}{c(c+1)} \cdot \frac{x^2}{2!} + \frac{a(a+1)(a+2).b(b+1)(b+2)}{c(c+1)(c+2)} \frac{x^2}{3!} + \ldots \quad \ldots(1)$$

where $-1 < x < 1$ and $c > a + b - 1$, is called Hypergeometric series.

In terms of Factorial Powers, this can be re-written as

$$F\left(\begin{matrix} a, b \\ c \end{matrix} \middle| x \right) = 1 + \sum_{j=1}^{\infty} \frac{a^{[j]}\, b^{[j]}\, x^j}{c^{[j]} . j!} = 1 + \sum_{j=1}^{\infty} \frac{(-a)^{(j)}\, (-b)^{(j)}}{(-c)^{(j)} . j!} (-x)^j \quad \ldots(2)$$

as $a^{[j]} = (-1)^j\, (-a)^{(j)}$, etc.

Sometimes, the L.H.S. of (1) written $F(a, b\ ;\ c \mid x)$. Some special cases are :

$$F(-n, 1; -1; -x) = (1+x)^n ; \quad F(1, p; p : x) = (1-x)^{-1}$$

$$F(1,\ 1; 2; -x) = (1/x) \ln(1+x) ; F(\tfrac{1}{2}, \tfrac{1}{2}; \tfrac{3}{2}; x^2) = (1/x) \sin^{-1} x.$$

The H.G. series (1) are the solutions of the differential equations

$$t(1-t)(d^2x/dt^2) + [c - (1+a+b)\, t](dx/dt) - abx = 0. \quad \ldots(3)$$

FOR STATISTICAL THEORY

Lebesgue's Theorem $\mathsf{E}\left(\lim_{n \uparrow \infty} X_n \right) = \lim_{n \uparrow \infty} \mathsf{E}(X_n)$

Differentiation Under the Expectation Sign (DUES)

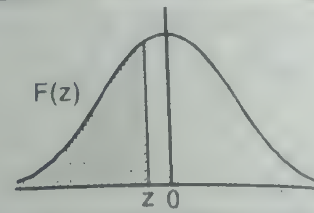
$$\frac{d}{dt} \mathsf{E}(X_t) = \mathsf{E}\left(\frac{d}{dt} X_t \right), \quad a < t < b.$$

Cinema is the most beautiful fraud in the world. (Gean-L Goduray)
But Music is the strongest form of magic. (Marilyn Manson)

Statistical Tables

Table 1. Standard Normal Distribution Function

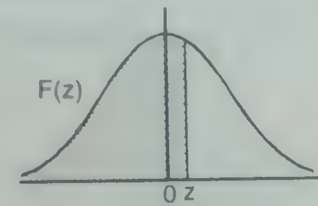
$$P(Z \leq z) = F(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-t^2/2}\, dt$$



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
–5.0	0.0000003									
–4.0	0.00003									
–3.5	0.0002									
–3.4	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002
–3.3	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003
–3.2	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	0.0005
–3.1	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007
–3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
–2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
–2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
–2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
–2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
–2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
–2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
–2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
–2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
–2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
–2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
–1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
–1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
–1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
–1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
–1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
–1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
–1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
–1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
–1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
–1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
–0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
–0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
–0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
–0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
–0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
–0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
–0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
–0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
–0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
–0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641

Table 1. Standard Normal Distribution Function

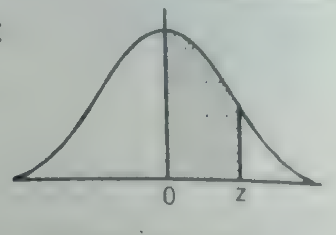
$$P\{Z \leq z\} = F(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-t^2/2}\, dt$$



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998									
4.0	0.99997									
5.0	0.9999997									

Table 2. Areas under the Standard Normal Curve from 0 to z

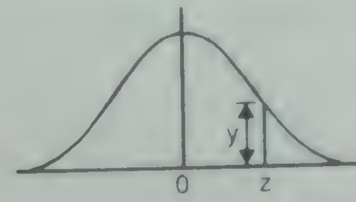
$$P\{0 \le Z \le z\} = \psi(z) = \int_0^z \frac{1}{\sqrt{2\pi}} e^{-t^2/2}\, dt$$



z	0	1	2	3	4	5	6	7	8	9
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0754
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2258	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2518	.2549
0.7	.2580	.2612	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2996	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990
3.1	.4990	.4991	.4991	.4991	.4992	.4992	.4992	.4992	.4993	.4993
3.2	.4993	.4993	.4994	.4994	.4994	.4994	.4994	.4995	.4995	.4995
3.3	.4995	.4995	.4995	.4996	.4996	.4996	.4996	.4996	.4996	.4997
3.4	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4998
3.5	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998
3.6	.4998	.4998	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999
3.7	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999
3.8	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999
3.9	.5000	.5000	.5000	.5000	.5000	.5000	.5000	.5000	.5000	.5000

Table 3. Ordinates (y) of the Standard Normal Curve at z

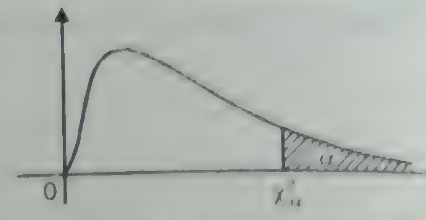
$$y = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$$



z	0	1	2	3	4	5	6	7	8	9
0.0	.3989	.3989	.3989	.3988	.3986	.3984	.3982	.3980	.3977	.3973
0.1	.3970	.3965	.3961	.3956	.3951	.3945	.3939	.3932	.3925	.3918
0.2	.3910	.3902	.3894	.3885	.3876	.3867	.3857	.3847	.3836	.3825
0.3	.3814	.3802	.3790	.3778	.3765	.3752	.3739	.3725	.3712	.3697
0.4	.3683	.3668	.3653	.3637	.3621	.3605	.3589	.3572	.3555	.3538
0.5	.3521	.3503	.3485	.3467	.3448	.3429	.3410	.3391	.3372	.3352
0.6	.3332	.3312	.3292	.3271	.3251	.3230	.3209	.3187	.3166	.3144
0.7	.3123	.3101	.3079	.3056	.3034	.3011	.2989	.2966	.2943	.2920
0.8	.2897	.2874	.2850	.2827	.2803	.2780	.2756	.2732	.2709	.2685
0.9	.2661	.2637	.2613	.2589	.2565	.2541	.2516	.2492	.2468	.2444
1.0	.2420	.2396	.2371	.2347	.2323	.2299	.2275	.2251	.2227	.2203
1.1	.2179	.2155	.2131	.2107	.2083	.2059	.2036	.2012	.1989	.1965
1.2	.1942	.1919	.1895	.1872	.1849	.1826	.1804	.1781	.1758	.1736
1.3	.1714	.1691	.1669	.1647	.1626	.1604	.1582	.1561	.1539	.1518
1.4	.1497	.1476	.1456	.1435	.1415	.1394	.1374	.1354	.1334	.1315
1.5	.1295	.1276	.1257	.1238	.1219	.1200	.1182	.1163	.1145	.1127
1.6	.1109	.1092	.1074	.1057	.1040	.1023	.1006	.0989	.0973	.0957
1.7	.0940	.0925	.0909	.0893	.0878	.0863	.0848	.0833	.0818	.0804
1.8	.0790	.0775	.0761	.0748	.0734	.0721	.0707	.0694	.0681	.0669
1.9	.0656	.0644	.0632	.0620	.0608	.0596	.0584	.0573	.0562	.0551
2.0	.0540	.0529	.0519	.0508	.0498	.0488	.0478	.0468	.0459	.0449
2.1	.0440	.0431	.0422	.0413	.0404	.0396	.0387	.0379	.0371	.0363
2.2	.0355	.0347	.0339	.0332	.0325	.0317	.0310	.0303	.0297	.0290
2.3	.0283	.0277	.0270	.0264	.0258	.0252	.0246	.0241	.0235	.0229
2.4	.0224	.0219	.0213	.0208	.0203	.0198	.0194	.0189	.0184	.0180
2.5	.0175	.0171	.0167	.0163	.0158	.0154	.0151	.0147	.0143	.0139
2.6	.0136	.0132	.0129	.0126	.0122	.0119	.0116	.0113	.0110	.0107
2.7	.0104	.0101	.0099	.0096	.0093	.0091	.0088	.0086	.0084	.0081
2.8	.0079	.0077	.0075	.0073	.0071	.0069	.0067	.0065	.0063	.0061
2.9	.0060	.0058	.0056	.0055	.0053	.0051	.0050	.0048	.0047	.0046
3.0	.0044	.0043	.0042	.0040	.0039	.0038	.0037	.0036	.0035	.0034
3.1	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026	.0025	.0025
3.2	.0024	.0023	.0022	.0022	.0021	.0020	.0020	.0019	.0018	.0018
3.3	.0017	.0017	.0016	.0016	.0015	.0015	.0014	.0014	.0013	.0013
3.4	.0012	.0012	.0012	.0011	.0011	.0010	.0010	.0010	.0009	.0009
3.5	.0009	.0008	.0008	.0008	.0008	.0007	.0007	.0007	.0007	.0006
3.6	.0006	.0006	.0006	.0005	.0005	.0005	.0005	.0005	.0005	.0004
3.7	.0004	.0004	.0004	.0004	.0004	.0004	.0003	.0003	.0003	.0003
3.8	.0003	.0003	.0003	.0003	.0003	.0002	.0002	.0002	.0002	.0002
3.9	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0001	.0001

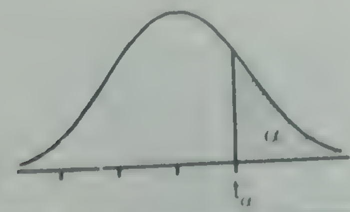
Table 4. Values of $\chi^2_\alpha(\nu)$

$\nu = n - 1$



ν	$\alpha = 0.995$	$\alpha = 0.99$	$\alpha = 0.975$	$\alpha = 0.95$	$\alpha = 0.05$	$\alpha = 0.025$	$\alpha = 0.01$	$\alpha = 0.005$	ν
1	0.0000393	0.000157	0.000982	0.00393	3.841	5.024	6.635	7.879	1
2	0.0100	0.0201	0.0506	0.103	5.991	7.378	9.210	10.597	2
3	0.0717	0.115	0.216	0.352	7.815	9.348	11.345	12.838	3
4	0.207	0.297	0.484	0.711	9.488	11.143	13.277	14.860	4
5	0.412	0.554	0.831	1.145	11.070	12.832	15.056	16.750	5
6	0.676	0.872	1.237	1.635	12.592	14.449	16.812	18.548	6
7	0.989	1.239	1.690	2.167	14.067	16.013	18.475	20.278	7
8	1.344	1.646	2.180	2.733	15.507	17.535	20.090	21.955	8
9	1.735	2.088	2.700	3.325	16.919	19.023	21.666	23.589	9
10	2.156	2.558	3.247	3.940	18.307	20.483	23.209	25.188	10
11	2.603	3.053	3.816	4.575	19.675	21.920	24.725	26.757	11
12	3.074	3.571	4.404	5.226	21.026	23.337	26.217	28.300	12
13	3.565	4.107	5.009	5.892	22.362	24.736	27.688	29.819	13
14	4.075	4.660	5.629	6.571	23.685	26.119	29. 141	31.319	14
15	4.601	5.229	6.262	7.261	24.996	27.488	30.578	32.801	15
16	5.142	5.812	6.908	7.962	26.296	28.845	32.000	34.267	16
17	5.697	6.408	7.564	8.672	27.587	30.191	33.409	35.718	17
18	6.265	7.015	8.231	9.390	28.869	31.526	34.805	37.156	18
19	6.844	7.633	8.907	10.117	30.144	32.852	36.191	38.582	19
20	7.434	8.260	9.591	10.851	31.410	34.170	37.566	39.997	20
21	8.034	8.897	10.283	11.591	32.671	35.479	38.932	41.401	21
22	8.643	9.542	10.982	12.338	33.924	36.781	40.289	42.796	22
23	9.260	10.196	11.689	13.091	35.172	38.076	41.638	44.181	23
24	9.886	10.856	12.401	13.484	36.415	39.364	42.980	45.558	24
25	10.520	11.524	13.120	14.611	37.652	40.646	44.314	46.928	25
26	11.160	12.198	13.844	15.379	38.885	41.923	45.642	48.290	26
27	11.808	12.879	14.573	16.151	40.113	43.194	46.963	49.645	27
28	12.461	13.565	15.308	16.928	41.337	44.461	48.278	50.993	28
29	13.121	14.256	16.047	17.708	42.557	45.772	49.588	52.336	29
30	13.787	14.953	16.791	18.493	43.773	46.979	50.892	53.672	30
40	20.706	22.164	24.433	26.509	55.758	59.342	63.691	66.766	40
50	27.991	29.707	32.357	34.764	67.505	71.420	76.154	79.490	50
60	35.535	37.485	40.482	43.118	79.082	83.298	88.379	91.952	60
70	43.275	45.442	48.758	51.739	90.531	95.023	100.425	104.215	70
80	51.172	53.540	57.153	60.391	101.879	106.629	112.329	116.321	80
90	59.196	61.754	65.646	69.126	113.145	118.136	124.116	128.299	90
100	67.328	70.065	74.222	77.929	124.342	129.561	135.807	140.169	100

Table 5. Values of $t_\alpha(\nu)$



ν	$\alpha = 0.10$	$\alpha = 0.05$	$\alpha = 0.025$	$\alpha = 0.01$	$\alpha = 0.00833$	$\alpha = 0.00625$	$\alpha = 0.005$	ν
1	3.078	6.314	12.706	31.821	38.204	50.923	63.657	1
2	1.886	2.920	4.303	6.965	7.650	8.860	9.925	2
3	1.638	2.353	3.182	4.541	4.857	5.392	5.841	3
4	1.533	2.132	2.776	3.747	3.961	4.315	4.604	4
5	1.476	2.015	2.571	3.365	3.534	3.810	4.032	5
6	1.440	1.943	2.447	3.143	3.288	3.521	3.707	6
7	1.415	1.895	2.365	2.998	3.128	3.335	3.499	7
8	1.397	1.860	2.306	2.896	3.016	3.206	3.355	8
9	1.383	1.833	2.262	2.821	2.934	3.111	3.250	9
10	1.372	1.812	2.228	2.764	2.870	3.038	3.169	10
11	1.363	1.796	2.201	2.718	2.820	2.981	3.106	11
12	1.356	1.782	2.179	2.681	2.780	2.934	3.055	12
13	1.350	1.771	2.160	2.650	2.746	2.896	3.012	13
14	1.345	1.761	2.145	2.624	2.718	2.864	2.977	14
15	1.341	1.753	2.131	2.602	2.694	2.837	2.947	15
16	1.337	1.746	2.120	2.583	2.673	2.813	2.921	16
17	1.333	1.740	2.110	2.567	2.655	2.793	2.898	17
18	1.330	1.734	2.101	2.552	2.639	2.775	2.878	18
19	1.328	1.729	2.093	2.539	2.625	2.759	2.861	19
20	1.325	1.725	2.086	2.528	2.613	2.744	2.845	20
21	1.323	1.721	2.080	2.518	2.602	2.732	2.831	21
22	1.321	1.717	2.074	2.508	2.591	2.720	2.819	22
23	1.319	1.714	2.069	2.500	2.582	2.710	2.807	23
24	1.318	1.711	2.064	2.492	2.574	2.700	2.797	24
25	1.316	1.708	2.060	2.485	2.566	2.692	2.787	25
26	1.315	1.706	2.056	2.479	2.559	2.684	2.779	26
27	1.314	1.703	2.052	2.473	2.553	2.676	2.771	27
28	1.313	1.701	2.048	2.467	2.547	2.669	2.763	28
29	1.311	1.699	2.045	2.462	2.541	2.663	2.756	29
inf.	1.282	1.645	1.960	2.326	2.394	2.498	2.576	inf.

The less routine, the more life. (A. Bronson Alcot)

Brief Bibliography

1. ***A.M. Mood, F.A. Gray bill, D.C. Boes*** : *Introduction to the Theory of Statistics* : 3/e (1974), Mc-Graw Hill, N.Y.
2. ***Blake (Ian F.)*** : *An Introduction to Applied Probability* (1979), John Wiley & Sons, N.Y.
3. ***B.W. Lindgren*** : *Statistical Theory,* 3/e (1976), Macmillan Publishing Co. N.Y.
4. ***C.R. Heath Cote*** : *Probability : Elements of Mathematical Theory* (1971), George Allen & Unwin Ltd. London, U.K.
5. ***C.R. Rao*** : *Linear Statistical Inference and its applications* 2/e (1973), John Willey, N.Y.
6. ***David Stirzaker*** : *Elementary Probability* 2/e, 2003. Cambridge University Press, U.K.
7. ***E. Parzen*** : *Modern Probability Theory and its Applications* (1960), John Willey, N.Y.
8. ***Feldman & Fox*** : *Probability : The Mathematics of Uncertainty,* 1991, *Dekker.*
9. ***G.G. Roussas*** : *First Course in Mathematical Statistics* (1973), Addison Wesley Publishing Co. N.Y.
10. ***G.P. Beaumont*** : *Intermediate Mathematical Statistics* (1980), Chapman & Hall, London, U.K.
11. ***G.P. Beaumont*** : *Probability & Random Variables* (1986), Ellis Horwood Limited, England.
12. ***Harold Cramer*** : *Mathematical Methods of Statistics* (1946), Princeton Univ. Press, N.J.
13. ***H.D. Brunk*** : An Introduction to Mathematical Statistics 3/e, John Willey, N.Y.
14. ***H.J. Larson*** : *Introduction to Probability Theory and Statistical Inference* 3/e, (1982), John Willey, N.Y.
15. ***Hoel, Port, Stone*** : (i) *Introduction to Probability Theory* (1971), (ii) *Introduction to Statistical Theory* (1971), (iii) *Introduction to Stochastic Processes* (1972), Houghton Mifflin Company, Boston.
16. ***J. Aitchison*** : *Statistics* I (1968), *Statistics* II (1971), Oliver & Boyd, London.
17. ***J.V. Uspensky*** : *Introduction to Mathematical Probability* (1937), Mc-Graw Hill, N.Y.
18. ***John Freund*** : *Mathematical Statistics with Applications* 7/e, (Reprint : 2004) Pearson Education Inc. (India).

19. ***Marek Capinski & Tomasz Zastawniak*** : *Probability Through Problems* (Aug. 2000), Springer, N.Y.
20. ***Meyer Dwass*** : *Probability & Statistics* (1970), W.A. Benjamin, N.Y.
21. ***M. Kendall & A. Stuart*** : *The Advanced Theory of Statistics* Vol. 1, 5/e (1986), Vol. 2, 4/e (1979); G. Griffin & Co. London, U.K.
22. ***N.A. Rahman*** : *Exercises in Probability & Statistics* 2/e (1983), Griffin, London.
23. ***N.L. Johnson*** : *Continuous Multivariate Distributions*, Houghton Mifflin Co., Boston.
24. ***N.L. Johnson & Samuel Kotz*** : *Distributions in Statistics.* (i) *Discrete Distribution* (1972), (ii) *Continuous Univariate Distributions* Vol.1 (1973), Vol.2 (1973), Houghton Mifflin Company, Boston.
25. ***P.E. Pfeiffer*** : *Probability for Applications*, 1990, Springer-Verlag.
26. ***Paul L. Meyer*** : *Introductory Probability and Statistical Applications*, 2/e (1979), Addison Wesley, N.Y.
27. ***Robert Ash*** : *Basic Probability Theory* (1970), John Wiley N.Y.
28. ***Robert Ash*** : *Probability & Measure Theory* 2/e, (Reprint : 2005), Elsevier Inc. New Delhi (India).
29. ***R.V. Hogg & A.T. Craig*** : *Introduction to Mathematical Statistics*, 5/e, (Reprint : 2002), Pearson Education Inc. (India).
30. ***Sanjay Arora & Bansi Lal*** : *New Mathematical Statistics*, (1989), Satya Prakashan, N. Delhi.
31. ***Sheldon M. Ross*** : *A First Course in Probability* 6/e, (Reprint : 2002), Pearson Inc. (India).
32. ***Sheldon M. Ross*** : *Introduction to Probability Models* 8/e, (Reprint : 2003), Elsevier Inc. New Delhi.
33. ***Sheldon M. Ross*** : *Introduction to Probability and Statistics for Engineers & Scientists* 3/e; (Reprint : 2004), Elsevier Inc, New Delhi.
34. ***S.S. Wilks*** : *Mathematical Statistics* (1962), John Wiley, N.Y.
35. ***T. Cacoullos*** : *Exercises in Probability* (Reprint : 1993) Narosa Publishing House New Delhi.
36. ***Tom. M. Apostol*** : *Calculus* Vol. II, 2/e (1960), John Wiley, N.Y. *Mathematical Analysis* 2/e (1974), Addison-Wesley, N.Y.
37. ***V.K. Rohtagi*** : *An Introduction to Probability Theory and Mathematical Statistics* (1976), John Wiley, N.Y.
38. ***V.K. Rohatgi*** : *Statistical Inference* 1984, John Wiley & Sons, N.Y.
39. ***Walpole, Myers, Myers, Ye*** : *Probability & Statistics for Engineers & Scientists*, 7/e, (Reprint : 2001) Pearson Educations Inc. (India).
40. ***W. Feller*** : *An Introduction to Probability Theory and its Applications* : Vol.I, 3/e (1968); Vol II (1966), John Wiley, N.Y.

A bone to the dog is not charity. Charity is the bone shared with the dog, when you are just as hungry as the dog. (Jack London)

The secret to creativity is knowing how to hide your sources.
(Albert Einstein)

Subject Index

D

Probability begins and ends with probability. (J.M. Keynes)

Some Discrete Distributions

Name-Notation	pmf : $f(x)$	mgf	$E(X)$	$Var(X)$
bin (n, p)	$\binom{n}{x} q^{n-x} p^x, x = 0, 1, \ldots, n$	$(q + pe^t)^n$	np	npq
Pois (λ)	$e^{-\lambda} \lambda^x / x!, \ x = 0, 1, 2, \ldots$	$e^{\lambda(e^t - 1)}$	λ	λ
goem (p)	$pq^x, x = 0, 1, 2, \ldots.$	$p/(1 - qe^t)$	q/p	q/p^2
gem (p)	$pq^{x-1}, x = 1, 2, \ldots$	$pe^t/(1 - qe^t)$	$1/p$	q/p^2
NB (k, p)	$\binom{x+k-1}{k-1} p^k q^x, \ x = 0, 1, \ldots$	$[p/(1 - qe^t)]^k$	kq/p	kq/p^2
NB* (k, p)	$\binom{y-1}{k-1} p^k q^{x-k}, \ y = k, k+1, \ldots$	$[pe^t/(1 - qe^t)]^k$	k/p	kq/p^2
H-G (N, M, n)	$\binom{M}{x}\binom{N-M}{n-x} \Big/ \binom{N}{n}$	Not useful	np	$\dfrac{npq\,(N-n)}{N-1}$
	$\max\{0, M + n - N\} \le x \le \min\{M, n\}.$		Here $p = M/N$	
$M(n, k, p)$	$\binom{n}{x_1, \ldots, x_k} p_1^{x_1} \ldots p_k^{x_k}$	$(p_1 e^{t_1} + \ldots + p_k e^{t_k})^n$		
	$0 < p_i < 1, \ \sum x_i = n, \ \sum p_i = 1. \ X_i \sim \text{bin}\,(n, p_i)$			

Some Continuous Distributions

Name-Notation	density : $f(x)$	mgf	$E(X)$	$Var(X)$
$N(\mu, \sigma^2)$	$(\sqrt{2\pi\sigma})^{-1} e^{-(x-\mu)^2/2\sigma^2} \quad -\infty < x < \infty$	$e^{\mu t + (\sigma^2 t^2/2)}$	$\mu,$	σ^2
ln $N(\mu, \sigma^2)$	$(x\sqrt{2\pi\sigma})^{-1} e^{-\frac{1}{2}[(\log x - \mu)/\sigma]^2}, \ x > 0$	does not exist	$\mu'_k = e^{\mu k + \sigma^2 k^2/2}$	
$U(a, b)$	$1/(b-a), a < x < b$	$(e^{bt} - e^{at})/t(b-a)$	$(a+b)/2$	$(b-a)^2/12$
Expo (λ)	$\lambda e^{-\lambda x}, \lambda > 0, x > 0$	$[1 - (t/\lambda)]^{-1}$	$1/\lambda$	$1/\lambda^2$
gam (a, λ)	$\dfrac{(\lambda e^{-\lambda x})(\lambda x)^{a-1}}{\Gamma(a)}, a > 0, \lambda > 0, x > 0$	$[1 - (t/\lambda)]^{-a}$	$\dfrac{a}{\lambda}$	$\dfrac{a}{\lambda^2}$
$\chi^2_{(n)}$	$\dfrac{e^{-x/2} x^{(n/2)-1}}{\Gamma(n/2) 2^{n/2}}, \ x > 0$	$(1-2t)^{-n/2}$	n	$2n$
$B_{\text{I}}(a, b)$	$\dfrac{x^{a-1}(1-x)^{b-1}}{B(a, b)}, a > 0, b > 0, 0 < x < 1$	Not useful	$\dfrac{a}{a+b},$	$\dfrac{ab}{(a+b)^2(a+b+1)}$
$B_{\text{II}}(a, b)$	$\dfrac{x^{a-1}}{B(a, b)(1+x)^{a+b}}, \ x > 0$	Not useful	$\dfrac{a}{b-1},$	$\dfrac{a(a+b-1)}{(b-1)^2(b-2)}$
Lap (μ, λ)	$\frac{1}{2}\lambda e^{-\lambda\lvert x-\mu\rvert}, \ \lambda > 0, \ \lvert x\rvert < \infty$	$\dfrac{\lambda^2 e^{\mu t}}{\lambda^2 - t^2}$	μ	$\dfrac{2}{\lambda^2}$
Chy (a, b)	$b/\pi[b^2 + (x-a)^2], \ b > 0, \ \lvert x\rvert < \infty$	$M(it) = e^{iat - b\lvert t\rvert}$	Not exists	
Weib (α, β)	$\alpha\beta x^{\beta-1} e^{-\alpha x^\beta}, \ x > 0, \ \alpha > 0, \beta > 0$		$\mu = \alpha^{-1/\beta}\Gamma(1+\beta^{-1}),$	$\sigma^2 = \alpha^{-2/\beta}\{\Gamma(1+2\beta^{-1}) - \Gamma^2(1+\beta^{-1})\}$

About the Book

Introducing Probability & Statistics 2/e is aimed at reaching the mid-high, although it starts at the lowest level of basic fundas. It intends to inculcate in the reader's mind an intuitive as well as the mathematical feel for the subject which is so essential for scientific thinking.

Most of the standard topics are included, but a few are omitted, which are given in authors' New Mathematical Statistics (1989). Just a glance over Partial Contents reveals the topics covered.

The vastly-covered topics are profusely illustrated by over 900 fully worked-out distinct Problems, of which 542 are starred and their solutions are provided at the back of the Text (pp : 659-818). This shall help rapid coverage and flexibility of available choices.

About the Authors

Dr. BANSI LAL

Bansi Lal had Honours in Mathematics from Panjab University in 1953 ; postgraduation in 1955 and subsequently passed ALL the special (optional) papers of Mathematics which Panjab University had listed in its syllabus during that span, joined Panjab University (C) College of Delhi in 1956 and later joined K.M. College in 1960. He has written extensively during times when things were rather scarce. NEW MATHEMATICAL STATISTICS (1989), ELEMENTARY HYDRODYNAMICS (1960), DIFFERENTIAL GEOMETRY (VECTORIAL APPROACH) 1961, Theoretical Hydro Dyanmics (1963), Mathematical Theory of Electromagnetism (1964). Many of these books are first to appear in India. Some under-graduate works include INTRODUCING METRIC SPACES, VECTOR ANALYSIS, VECTOR CALCULUS, REAL ANALYSIS, MATHEMATICS OF PROBABILITY and STATISTICS, TWO DIMENSIONAL COORDINATE GEOMETRY and PARTICLE DYNAMICS.

He has contributed extensively to Journals in America, West Germany and India. He was also visiting Prof. of Mathematics University of Iowa, Iowa City, U.S.A.

Dr. SANJAY ARORA

Sanjay Arora graduated from the University of Delhi in 1980 and M.Phil in 1981 with the usual record performance of a brilliant scholar. He joined the department of Mathematics, S.V. College (South Delhi) in 1981. In 1984, he joined the research activities in the Department of Mathematics and Statistics of Colorado State University, U.S.A. He has contributed valuable research work in Mathematical Statistics under the guidance of Prof. J.N. Srivastava, an international authority on Designs of Experiment. Sometimes, Sanjay Arora was with the Department of Mathematics and Statistics, Howard University, Washington D.C. U.S.A.

Dr. Sanjay has co-authored some books : INTRODUCING METRIC SPACE, NEW MATHEMATICAL STATISTICS, REAL ANALYSIS, VECTOR ANALYSIS, VECTOR CALCULUS and TWO-DIMENSIONAL ANALYTICAL GEOMETRY are some of his well-known books. His current interests is centred around DESIGN OF EXPERIMENTS.

SUDHA ARORA

Sudha Arora graduated in 1987 and post-graduated in 1989, both from the University of Madras, and completed M. Phil. (Mathematics) in 1992 from the University of Delhi. Right from schooling to the University level, she has had a consistent record of attaining far ahead of distinction credits in her academic achievements. She has taught Honours Courses in Mathematics to various classes commencing her teaching career in January 1992 in a Delhi-University College. She has also taught in Kirorimal College as well as Miranda House.

At present, she is a permanent member of the Department of Mathematics, Deen Dayal Upadhyaya College (University of Delhi) Delhi. Her research interest include Modern Abstract Algebra and Operation Research. She has co-authored : Mathematics Main for C.S.E. Standard Papers (1996) : a work first of its kind in India and Elementary Real Analysis (1996).

Satya Prakashan New Delhi

Phone +91-11-28715679
+91-11-28713653

Fax +91-11-28713593

E-mail deepakhanda@hotmail.com
deepak@satyaprakashan.com
ajay@satyaprakashan.com

Visit us www.satyaprakashan.com